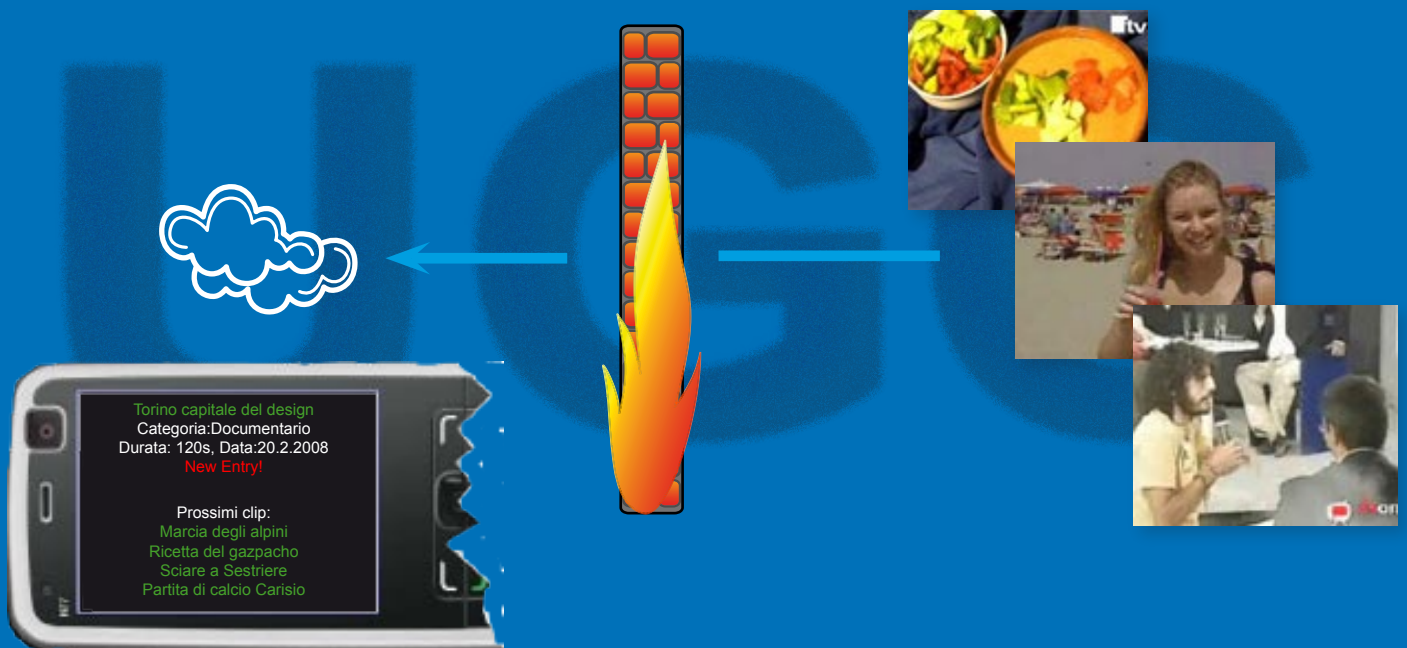
Elettronica e telecomunicazioni

Anno LVII
Numero 1
Aprile 2008

Editoriale

Sistema per la produzione automatica di servizi Audio e Video per Mobile TV e Mobile Web

Distorsioni nella stereoscopia tradizionale

Rilevazione, correzione o mascheramento degli errori

Edizione ottimizzata per la stampa.
La rivista è disponibile su web
alla URL www.crit.rai.it/eletel.htm

Indice

Anno LVII
N° 1
Aprile 2008

Rivista quadrimestrale a cura della Rai

Editoriale

ing. Gianfranco **Barbieri**
Direttore di
"Elettronica e
Telecomunicazioni"

La possibilità di erogare servizi televisivi in tempo reale su ricevitori palmari (battezzati ormai "videofonini" nel gergo della comunicazione), apparati caratterizzati da batterie di capacità limitata e destinati ad operare in condizioni di ricezione critiche, costituisce oggi l'ultima frontiera dell'evoluzione tecnologica che da più di un decennio sta radicalmente trasformando il mondo della radiodiffusione.

Servizi sperimentali di TV Mobile sono stati lanciati in una cinquantina di paesi mentre in altri, tra cui l'Italia, si è già alla fase di veri e proprii servizi commerciali. Nel marzo 2008 u.s. la Commissione Europea ha formalmente appoggiato il Sistema DVB-H come Standard raccomandato per la TV Mobile in Europa ed ha invitato gli stati membri ad incoraggiare lo sviluppo della relativa tecnologia.

Per i radiodiffusori, il cimentarsi nel mercato della TV Mobile rappresenta una sfida di non poco conto. Mentre, infatti, la principale innovazione offerta dalla trasmissione ai ricevitori palmari viene comunemente percepita come "La Televisione per utenti in mobilità", una molteplicità di servizi addizionali può, tuttavia, trarre giovamento dalle incrementate funzionalità delle tecnologie di radiodiffusione digitale. La stessa radiofonia, che, peraltro, da moltissimi anni si avvale della possibilità di funzionare su ricevitori mobili, può trarre vantaggio dall'impiego dei metadati per la scelta e la registrazione dei programmi e del ricco materiale multimediale che può essere associato al segnale musicale per migliorare l'esperienza auditiva.

La TV Mobile può essere considerata un servizio a sé stante ma può, più utilmente, sfruttare le funzionalità della telefonia mobile per implementare servizi realmente interattivi. I risultati di varie campagne di servizi sperimentali rivelano che, da parte dell'utenza, esiste un notevole interesse a che sia possibile avere accesso ad un vasto campionario di contenuti multimediali sui ricevitori palmari. Questo scenario comporta necessariamente la necessità di creare nuove tipologie di programmi e di riorganizzare l'utilizzo dei materiali d'archivio; i problemi maggiori per i radiodiffusori sono dovuti ai costi per la produzione di contenuti ad-hoc e per la realizzazione di una rete di diffusione a copertura nazionale cui si dovrebbe far fronte nel contesto di un mercato non ancora consolidato (ritorno sugli investimenti solo nel lungo termine). Risulta pertanto evidente l'esigenza di mettere a punto procedure che consentano di razionalizzare gli interventi sugli impianti e di ottimizzare i costi di produzione.

Editoriale

L'articolo "Sistema per la produzione automatica di servizi Audio e Video per Mobile TV e Mobile Web" descrive un sistema prototipale finalizzato alla creazione di servizi multimediali indirizzati ad utenti mobili, utilizzando materiali d'archivio selezionati ed elaborati automaticamente. Nello stesso articolo vengono descritti due servizi sperimentali basati sul medesimo sistema, operativi nell'area di Torino, rivolti sia a piattaforme di TV Mobile (DVB-H oppure T-DMB indifferentemente) sia al Web Mobile, tecnologie che convivono nei cellulari dell'ultima generazione.

Prosegue anche in questo numero la serie di contributi sulla TV Stereoscopica, tematica sulla quale il Centro Ricerche e Innovazione Tecnologica della Rai è impegnato in sintonia con gli studi che attualmente vengono effettuati presso i Laboratori di vari paesi.

L'articolo "Distorsioni nella stereoscopia tradizionale" riferisce su alcune problematiche legate alle distorsioni che si generano lungo la filiera di produzione, dovute in parte alla configurazione geometrica del sistema di ripresa, altre alle caratteristiche delle ottiche, altre ancora ai dispositivi di visualizzazione. Nel corso dell'articolo viene fatta una esauriente disamina teorica dei fattori distorcenti e vengono suggerite alcune soluzioni per la minimizzazione degli inconvenienti; viene altresì osservato come certi problemi causati dal processo fisiologico della visione umana non possono essere eliminati totalmente.

La rivelazione e susseguente correzione degli errori nei sistemi di trasmissione costituisce una delle tematiche più interessanti ed impegnative nell'attuale panorama delle telecomunicazioni. La rivoluzione digitale mette a disposizione tecnologie che consentono di sviluppare sistemi sempre più complessi e, contestualmente, i nuovi servizi vengono erogati in condizioni operative che diventano ogni giorno più sfidanti. La degradazione introdotta da un canale rumoroso sui segnali digitali sottoposti, in particolare, a processi di forte compressione renderebbe questi ultimi inutilizzabili in fase di decodifica se non venissero adottate sofisticate strategie di protezione; la codifica di canale assume pertanto un'importanza fondamentale nel consentire il corretto funzionamento dei nuovi servizi.

Per la serie "Cos'è, come funziona" viene pubblicata in questo numero la prima di una serie di schede tramite le quali gli autori intendono fornire informazioni di base sulle tecniche e sugli algoritmi utilizzati per la protezione dagli errori introdotti dal canale di diffusione o registrazione ed è focalizzata prevalentemente sui codici a blocco; uno di tali codici, che ha trovato ampio impiego nel campo delle telecomunicazioni, ed in particolare per i sistemi per la trasmissione e diffusione della TV digitale DVB e ATSC, è quello denominato RS (dal nome di Irving Reed e Gustave Solomon che lo proposero nel 1960).

Sistema per la produzione automatica di servizi Audio e Video per Mobile TV e Mobile Web

ing. Gino **Alberico**,
ing. Paolo **Casagranda**,
ing. Alberto **Messina**,
ing. Francesco **Russo**

Rai
Centro Ricerche e
InnovazioneTecnologica
Torino

Sommario

Questo articolo descrive un innovativo sistema prototipale sviluppato per creare servizi multimediali indirizzati ad utenti mobili, partendo da contenuti d'archivio selezionati ed elaborati automaticamente o da User Generated Contents. Il sistema si rivolge sia alla TV Mobile su DVB-H sia al Web Mobile, tecnologie che convivono nei cellulari di ultima generazione. L'articolo descrive inoltre due servizi sperimentali pilota basati sul sistema prototipale, entrambi disponibili nell'area di Torino. Il primo servizio consiste nella creazione automatica di un canale tematico di notizie che utilizza contenuti estratti dai telegiornali regionali, il secondo permette di distribuire User Generated Contents.

1. Introduzione

Nel corso degli ultimi anni, in molti paesi europei, sono stati lanciati diversi trial DVB-H e alcuni servizi commerciali di TV Mobile [1] da cui emerge, da parte dell'utenza, una richiesta di contenuti aggiuntivi e alternativi alla semplice riproposizione dei canali TV generalisti. Ad esempio sembrerebbero molto apprezzati canali tematici, oppure canali realizzati con contenuti dinamici di breve durata, fruibili nei brevi intervalli di tempo a disposizione durante l'uso in condizioni di mobilità [2,3]. I terminali/telefoni per la TV mobile non sono solo dei piccoli televisori ma dei "nuovi media" a tutti gli effetti, e, in linea con quanto M. McLuhan aveva previsto, necessitano di nuovi contenuti.

Inoltre, i telefoni cellulari utilizzati per la TV Mobile sono oramai dotati anche di diverse tecnologie wireless, quali GPRS, UMTS o Wi-Fi, attraverso le quali l'utente può attivare connessioni bi-direzionali a banda larga e collegarsi a portali ove può accedere a servizi "on-demand" o interattivi. Pertanto, grazie alla complementarietà delle reti broadcast DVB-H e delle reti mobili 3G (o Wi-Fi) bi-direzionali, si può realizzare un'infrastruttura di rete geograficamente e funzionalmente pervasiva adatta al cosiddetto "ambient media delivery".

Tuttavia, la costruzione di nuove infrastrutture di rete, basate su tecnologie sia di tipo broadcast (DVB-H) sia di rete mobile (3G), comporta elevati investimenti a fronte di ricavi incerti, in un mercato non ancora maturo e con un numero relativamente basso di clienti. Analogamente, le risorse economiche per la produzione di contenuti audiovisivi specifici per una nuova piattaforma di TV Mobile sono ingenti. Per tali motivi si sono esplorate le possibilità per fornire, a basso costo, contenuti aggiuntivi alla semplice ripetizione di canali TV già esistenti, attraverso lo sfruttamento di materiale d'archivio, riutilizzo di contenuti esistenti e utilizzo di contenuti generati da utenti (UGC, User Generated Contents). In particolare sono stati identificati due scenari applicativi in cui si realizzano due canali trasmessi nel pacchetto di servizi di TV mobile diffuso tramite DVB-H nell'area di Torino: un canale di news regionali (carousel) e un canale Video Blog basato su User Generated Contents.

Nell'articolo viene descritto un innovativo sistema prototipale per la creazione automatica, la pubblicazione e la messa in onda di canali "tematici" specifici per l'impiego su piattaforme di TV Mobile. Il sistema consente anche la pubblicazione dei contenuti su un sito "Mobile-Web", realizzando così un primo passo nell'integrazione di diversi "nuovi media". Si noti che nell'articolo si fa riferimento alla tecnologia DVB-H [4, 5, 6] in quanto, oltre ad essere oggetto della sperimentazione, è stata anche esplicitamente indicata dalla Commissione Europea come standard per la TV mobile. Tuttavia, i servizi e il sistema prototipale

Acronimi e sigle

DVB	Digital Video Broadcasting, consorzio nato in Europa per la creazione di standard globali televisivi e di trasmissione dati. Da qui nascono, tra gli altri, lo standard DVB-H per la TV Mobile e il DVB-T per la televisione ditale terrestre. (www.dvb.org)
DVB-H	Digital Video Broadcasting to Handheld, definito dal consorzio DVB e standardizzato nella specifica ETSI EN 302 304 e seguenti.
CBMS	Convergence of Broadcast and Mobile Services, sottogruppo del consorzio DVB che si occupa della creazione di specifiche e linee guida per la cooperazione tra piattaforme di telecomunicazioni diffusive e mobili
ESG	Electronic Service Guide, la guida elettronica ai servizi utilizzata da DVB-H, definita dal gruppo DVB CBMS.
FLUTE	File Delivery over Unidirectional Transport Protocol, descritto in RFC3926, protocollo per il trasporto di file su canale unidirezionale, scelto dal gruppo DVB TM-CBMS per la TV Mobile su DVB-H.
H.264	detto anche AVC (Advanced Video Coding) o MPEG-4 Part 10, descritto nel documento ISO/IEC 14496-10, + un formato di compressione video ad elevata efficienza.
MBMS	Multimedia Broadcast Multicast Services, protocollo all'interno dello standard UMTS per i servizi multicast e broadcast.
MJPEG	Motion JPEG, formato di compressione del video.
PCM	Pulse Code Modulation, formato digitale per la codifica dell'audio.
QVGA	Quarter VGA, risoluzione corrispondente a 320 x 240 pixel.
RTP	Real Time Transport Protocol, protocollo internet per il trasporto in tempo reale di video, audio e dati.
SDTV	Standard Definition Television, formato televisivo a definizione standard, tipico di sistemi televisivi analogici o digitali a 576 linee e simili.
UGC	User Generated Contents, i contenuti creati direttamente dagli utenti.
UMTS	Universal Mobile Telecommunications System, è lo standard di terza generazione per le telecomunicazioni mobili digitali.

descritti nel seguito sono ugualmente applicabili (con eventuali modifiche o limitazioni) anche ad altre piattaforme quali ad esempio il T-DMB, il sistema MediaFLO della società Qualcomm o la tecnologia UMTS-MBMS.

2. Contenuti

La Rai produce tre edizioni giornaliere per venti Telegiornali Regionali, incluse le rubriche meteo, con un rilevante impegno di risorse articolato su tutto il territorio nazionale. Ciascuna edizione è strutturata secondo uno schema che prevede dopo la sigla un primo sommario delle principali notizie della durata di circa 1 minuto, seguita dai servizi completi relativi alle varie notizie e in chiusura il meteo. Poiché, nel caso della TV mobile, la durata di circa 20 minuti dell'edizione completa viene ritenuta spesso eccessiva dagli utenti, si è partiti dall'idea di utilizzare il solo sommario iniziale dei TG-R realizzando un canale con tutti i sommari delle 20 regioni ripetuti in sequenza (modalità carousel) in un blocco della durata di circa 20-25 minuti (comprendendo opportuni clip separatori tra una regione e l'altra). In questo modo, mentre sui canali tradizionali la fruibilità territoriale di ciascun TGR è attualmente limitata su base regionale, nel servizio TV mobile l'utente avrebbe a disposizione ogni mezz'ora un sommario con le notizie della propria regione aggiornato all'ultima edizione disponibile. Così facendo è possibile anche valorizzare un potenziale interesse di pubblico presumibilmente assai vasto, legato da un lato ai flussi migratori che hanno caratterizzato il nostro paese nel recente passato e dall'altro alla mobilità di fine settimana e dei periodi di vacanza.

Inoltre, tramite connessione a larga banda su rete mobile, dal telefono DVB-H è possibile accedere a una pagina web con maggiori approfondimenti, associata al servizio di News Regionali.

Un'altra tipologia di contenuti attualmente molto popolari e utilizzati nella sperimentazione di nuovi servizi riguarda i contenuti generati dagli utenti, i cosiddetti UGC. Ci sono ormai diversi siti Web che mettono a disposizione spazio e strumenti per la pubblicazione e la ricerca di tali contenuti UGC, un esempio per tutti come riferimento è il sito di YouTube. L'ipotesi del canale UGC Video Blog per la TV mobile nasce dall'idea che i video clip più popolari fruiti solitamente in modalità "on-demand" hanno di solito un numero di accessi tale da giustificarne la trasmissione in un canale broadcast come quello DVB-H. Il sistema prototipale si basa sul presupposto di avere già un'infrastruttura per la raccolta e la pubblicazione su Web di contenuti UGC con i relativi metadati: in particolare durante la sperimentazione ci si è avvalsi di una collaborazione con l'agenzia AdnKronos (sezione del sito "Secondo te"). Dal repository Web si selezionano manualmente, e questo è l'unico ma importante intervento "umano" nel processo, i video-clip ritenuti validi e interessanti e con opportune transcodifiche e adattamenti si procede alla schedulazione/trasmissione su DVB-H nonché alla pubblicazione su Mobile-Web.

Entrambi i servizi di News Regionali (carousel) e UGC Video Blog sono in onda dall'area di Torino da giugno-settembre 2007.

3. Descrizione del sistema

Il sistema prototipale descritto nel seguito consente di creare servizi multimediali indirizzati ad utenti mobili in modo quasi automatico, partendo da UGC o contenuti di archivio selezionati e trasformati automaticamente. Questo sistema è stato ideato sia per servizi diffusivi di Mobile TV su DVB-H sia per Mobile Web interattivo su reti 3G. Entrambi i servizi sono indirizzati a dispositivi handheld. L'intervento umano è necessario solo in operazioni di alto livello come la validazione dei contenuti e il controllo sulla qualità.

Il sistema è flessibile nei suoi componenti centrali, in quanto è in grado di rivolgersi ad un'ampia varietà di servizi indirizzati alla diffusione su Mobile TV e alla pubblicazione sul web.

Il funzionamento di base del sistema prevede la

messa in onda sequenziale di una serie di videoclip, in una modalità chiamata carousel (figura 1). Il ciclo di trasmissione prosegue indefinitamente, con aggiornamenti continui.

Fig. 1 – Ciclo di trasmissione dei videoclip.

Due tipi di servizio A/V sono stati creati e dimostrati; il primo consiste nella generazione di un carousel di videoclip di news, con contenuti d'archivio, il secondo nella creazione di carousel usando videoclip fornite dagli utenti (UGC).

Il sistema è stato implementato usando componenti con le seguenti funzionalità:

- Cattura e taglio automatici delle news da archivio (in questo modo vengono generati i videoclip sorgenti)
- Raccolta dei metadati relativi ai videoclip; i metadati possono essere forniti dal broadcaster (nel caso delle News Regionali) o contenute in file XML forniti dagli utenti (nel caso UGC)
- Adattamento del formato dei videoclip sorgenti alla trasmissione sul canale di Mobile TV. E' inclusa la transcodifica del videoclip da alto bitrate e alta risoluzione ad un formato compatibile con i ricevitori utilizzati
- Adattamento del formato del videoclip sorgente per la pubblicazione su web
- Generazione di metadati in formato adatto alla trasmissione e pubblicazione sul web, e creazione dei Title Clip per ogni videoclip sorgente
- Schedulazione dei videoclip per la trasmissione
- Trasmissione e pubblicazione del servizio

Le funzioni sopra descritte necessitano chiaramente di una supervisione, che validerà i clip rispetto ai requisiti editoriali del Broadcaster o del Service Provider. La funzionalità di Cattura e taglio automatici nel caso degli UGC non è presente.

L'architettura generale del sistema è illustrata in figura 2.

Il blocco "Videoclip sorgenti" dipende dallo scenario implementato: UGC o News Regionali nel caso dei Servizi Pilota.

Lo scenario UGC utilizza il web per veicolare i clip generati dagli utenti. La seguente figura chiarisce lo scenario UGC (figura 3):

Nello scenario UGC, i videoclip non necessitano di cattura e taglio automatici, in quanto sono caricati direttamente da un sito ftp.

Nello scenario News Regionali, il componente Grabber si occupa della cattura e taglio automatici. Le funzioni del Grabber sono due. In primo luogo, estrae segmenti di contenuti d'archivio pre-registrati che contengono le News da elaborare. L'estrazione avviene per ognuna delle 20 sedi Regionali Rai, attraverso la rete intranet. In secondo luogo, il Grabber taglia in modo preciso

3GPP H264/AAC Encoder
- Trascodifica
- Titolazione

Streaming Server
Generazione del carousel

Sistema DVB-H

MUX
Modulazione
Trasmissione

Rete DVB-H

Content Management
Scheduling & Monitoring

Rete IP

Generazione dei
Title Clip

Amministrazione
Remota

Repository per
Videoclip sorgenti

Rete IP

Videoclip sorgenti
(Lanci delle News
Regionali o UGC)

Torino capitale del design
Categoria:Documentario
Durata: 120s, Data:20.2.2008
New Entry!

Prossimi clip:
Marcia degli alpini
Ricetta del gazpacho
Sciare a Sestriere
Partita di calcio Carisio

Title Clip animato

Fig. 2 – Schema architetturale del sistema.

il segmento di contenuti A/V, in modo tale che il clip inizi esattamente quando inizia il TGR, e finisca esattamente alla fine dei lanci delle notizie. La figura 4 illustra schematicamente l'estrazione delle notizie:

Nel processo di taglio delle News Regionali, un punto critico è la qualità dei clip sorgenti, che non deve solo essere sufficiente per l'adattamento al mezzo di pubblicazione: deve anche consentire un preciso riconoscimento automatico.

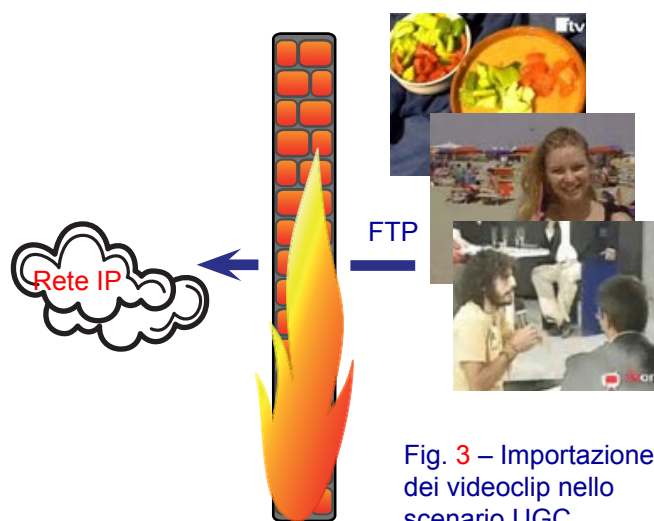



Fig. 3 – Importazione dei videoclip nello scenario UGC.

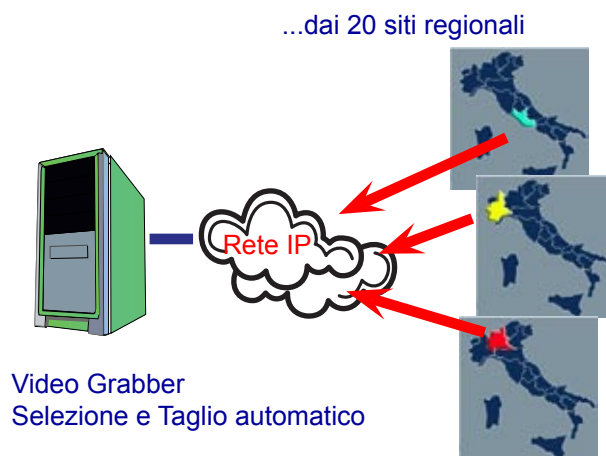



Fig. 4 – Importazione dei videoclip nello scenario News Regionali.

Una volta caricati e tagliati, i clip sorgenti devono essere codificati nel formato finale (adatto al mezzo di fruizione). Una scelta di progetto è stata l'utilizzo di un formato intermedio per l'archiviazione dei contenuti. Scegliendo opportunamente i parametri di codifica del formato intermedio, la grandezza del clip diminuisce, richiedendo meno spazio di archivio e tempi di trasferimento su rete più rapidi. Questa scelta influisce solo sullo scenario News Regionali, in quanto gli UGC arrivano generalmente con una qualità inferiore. Dopo alcuni test, è stato scelto il formato intermedio SDTV MJPEG a 4Mbps, con audio stereo PCM a 16 bit e 44 kHz, ad un bitrate di 1.5 Mbps, che consente un'ottima riduzione al formato per la TV Mobile (di solito QVGA H.264). L'adattamento al Mobile Web è più semplice al momento, in quanto un maggior numero di formati è adatto ai dispositivi finali; d'altra parte la grandezza del clip deve essere limitata per accelerare la velocità di download (e limitarne il costo): per questo è stato adottato un bitrate inferiore a quello scelto per la TV Mobile.

Per il mezzo TV Mobile, le informazioni sui videoclip trasmessi possono essere inviate agli utenti in due modi. Il primo consiste nell'uso della Guida Elettronica ai Servizi (DVB CBMS ESG, [7]), che consente un aggiornamento dinamico della descrizione dei contenuti trasmessi. Tuttavia la maggior parte dei terminali disponibile durante la presente ricerca non consentiva un veloce aggiornamento dinamico della ESG. Il secondo modo, qui considerato complementare e non alternativo al primo, consiste nell'uso di Title clip trasmessi prima del clip al quale si riferiscono; il Title clip presenta una lista scorrevole con metadati rilevanti come titolo, categoria, durata e informazioni sui clip successivi.

I metadati per la parte Mobile Web includono un piccolo frame del videoclip catturata automaticamente, insieme a tutte le informazioni utili su titolo, categoria, durata. Inoltre, l'interattività intrinseca del web permette le funzionalità aggiuntive di votazione del clip e ricerca nell'archivio dei clip.

La natura broadcast della TV Mobile rende invece necessario predisporre la schedulazione e messa in onda dei clip. Il prototipo usa un componente realizzato ad hoc, lo Schedulatore, che permette la schedulazione flessibile dei clip, come descritto nei paragrafi successivi. Lo Schedulatore controlla gli streaming server permettendo lo stream continuo dei contenuti e l'aggiunta dinamica di nuovi clip, senza intervento umano né pause durante il servizio. La piattaforma DVB-H raccoglie quindi i pacchetti RTP e la ESG su FLUTE [8] e li incapsula nella trama DVB-H, pronti per la distribuzione ai siti di trasmissione.

Nel caso del Mobile Web, l'ultimo passo del sistema consiste nella pubblicazione su web dei clip e relativi metadati

4. Generazione automatica dei clip dagli archivi A/V

Nell'applicazione dedicata alle News Regionali, il sistema utilizza un algoritmo basato sull'analisi del contenuto audiovisivo per selezionare le clip da pubblicare. L'algoritmo sfrutta una tecnica di corrispondenza tra elementi video per identificare l'inizio esatto del programma ed una tecnica di tracciamento di caratteristiche visuali basata sull'analisi della componente di luminanza e di saturazione per identificare la fine della sezione del programma dedicata ai titoli di testa.

La tecnica di corrispondenza tra elementi video utilizza un semplice processo di aggregazione nel quale gli elementi video (inquadrature) da analizzare sono associati a schemi di inquadrature conosciuti, usando il paradigma detto del "vicino più prossimo". In pratica, un certo elemento X è classificato come un'istanza di un elemento conosciuto A se il processo di aggregazione determina che il suo vettore caratteristico x è più simile al vettore caratteristico a di A che non a qualsiasi altro vettore. Le caratteristiche estratte per questo processo sono gli istogrammi di colore, le firme di texture di Tamura [9], e parametri di attività temporale quali la differenza

pixel a pixel della componente di luminanza di immagini consecutive.

La tecnica di tracciamento sfrutta alcune conoscenze del particolare dominio applicativo per conseguire con successo il suo compito. In particolare, essa sfrutta la peculiare configurazione di layout grafico dei programmi News Regionali, che è costituita da un primo sommario nel quale alcuni spezzoni dei servizi che saranno presentati in seguito sono presentati in sequenza, racchiusi da un layout fisso di tipo letter-box. Il contributo da studio non è caratterizzato da questa configurazione, quindi tracciare gli elementi caratteristici di quest'ultima è equivalente ad identificare i confini del sommario, che costituisce proprio il contenuto che si vuole pubblicare.

In concreto, questo è fatto analizzando l'intensità luminosa di alcune aree campione dell'immagine (figura 5), e scatenando un taglio del contenuto in corrispondenza all'istante nel quale questa caratteristica scende al di sotto di una soglia fissata. Questa soglia è stata determinata empiricamente su un insieme di dati campione utilizzati per l'addestramento, con l'obiettivo di ottimizzare l'accuratezza del rilevamento.

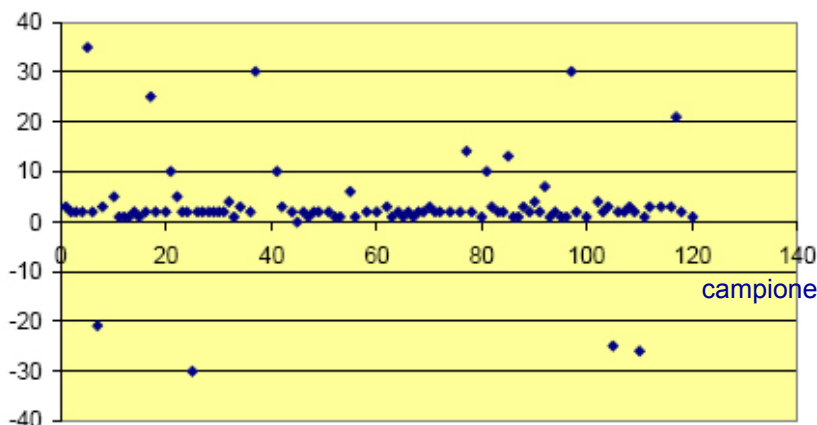
Per misurare le prestazioni dell'algoritmo sviluppato, si sono fatti alcuni test su un insieme di dati di riferimento costituito da N clip, e definendo il seguente indicatore di qualità:

$$q = \frac{1}{N_+ + N_- + N_{ok}} \left(N_{ok} + \sum_{i=1}^{N_+} \frac{d_i}{d'_i} + \sum_{i=1}^{N_-} \frac{d'_i}{d_i} \right) \quad (1)$$

Fig. 5 – Aree campione sottoposte a tracciamento.

dove N_+ è il numero di clip contenenti materiale in eccesso (falsi positivi), N_- è il numero di clip con materiale mancante (falsi negativi), N_{ok} è il numero di clip rilevate correttamente. Ovviamente $N = N_+ + N_- + N_{ok}$. Inoltre, d_i è la vera durata della clip, d'_i è la durata rilevata dall'algoritmo. Il parametro q varia tra 0 e 1, e associa valori più bassi a rilevamenti di minor qualità.

L'errore totale medio di rilevamento è stato misurato in 2.8 secondi (con 8.5 secondi di deviazione standard). La figura 6 riporta la distribuzione dell'errore di rilevamento sull'insieme di dati di test. In ipotesi di distribuzione gaussiana per l'errore,

Fig. 6 – Distribuzione dell'errore di rilevamento.

si è stimato che la probabilità che un punto di taglio rilevato sia più distante di 15 secondi dal punto effettivo è di 0.076, corrispondente ad un indice di qualità q=0.83. La prestazione di un tagliatore che operi deterministicamente in corrispondenza della durata media osservata sullo stesso campione, si attesta attorno ad un indice di qualità pari a q=0.64. Questo prova l'efficacia e l'utilità della tecnica implementata, la quale si è anche dimostrata parimenti efficace nella fase di operatività del servizio.

5. Clip di titolazione per TV Mobile

Servizi A/V tematici della Mobile TV simili al Video Blog sono strutturati in modo da mandare in onda sequenzialmente i videoclip, nella modalità definita carousel. Il Content Provider (o il Service Provider) dovrà quindi creare uno stacco tra i videoclip. Quest'ultima avrà diverse funzioni, oltre alla separazione vera e propria: rifletterà lo stile editoriale del servizio, ad esempio con un logo del servizio o della rete, colori e una colonna sonora adatta; descriverà inoltre i videoclip che seguono, fornendo un punto di orientamento utile agli utenti.

Nel servizio Video Blog, i Title Clip sono creati automaticamente. Una pagina con il logo aziendale precede un elenco in movimento di informazioni sul clip: titolo, categoria, durata, data di creazione. Viene segnalato se si tratta di un clip mandato in onda per la prima volta (quindi di interesse particolare). Segue una lista dei clip che verranno mandati in onda successivamente (figura 7) La durata del Title Clip è di circa 10 secondi.

Durante il test dei servizi pilota, il Title Clip si è rivelato molto utile, presentando un veloce accesso alle informazioni di palinsesto.

Notiamo infine che l'informazione contenuta nel Title Clip può anche essere trasportata dalla Electronic Service Guide del DVB-H (definita nelle specifiche del DVB-H, [7]).

6. Schedulatore

Il sistema di produzione automatica di servizi A/V si basa su uno Schedulatore progettato ad hoc. Lo Schedulatore permette la pianificazione flessibile dello streaming di insiemi di clip, attraverso l'utilizzo di template. I servizi pilota descritti in precedenza sono inclusi come casi particolari, e più in generale sono stati previste tre diverse modalità di funzionamento:

- Playlist Statica,
- Playlist Dinamica
- Video Blog

La modalità Playlist Statica consente di creare manualmente liste statiche di clip da trasmettere. Questa funzionalità è simile a quella di default dello streaming server Darwin, e consente di mandare in esecuzione sequenzialmente la playlist. Questa modalità è utilizzabile ad esempio quando i tempi di schedulazione delle clip da trasmettere siano identici ogni giorno (caso piuttosto raro in realtà).

La modalità Playlist Dinamica è invece indirizzata alla definizione di un palinsesto giornaliero o settimanale. E' possibile specificare i videoclip da inserire nel palinsesto con diverse opzioni di messa in onda:

- Modalità Sincrona: Consente di specificare l'esatto orario di messa in onda. In caso di due videoclip consecutivi schedulati

Fig. 7 – Esempio di Title Clip.

in Modalità Sincrona, in generale ci sarà un intervallo tra la fine del primo e l'inizio del secondo, in quanto non è possibile conoscere a priori la durata di ciascun videoclip. Per questo è necessario specificare un padding clip, da mandare in onda in questo intervallo. Il padding clip verrà ripetuto un numero di volte sufficiente per colmare l'intervallo, e dovrà essere abbastanza breve, in modo da non causare ritardi eccessivi nella messa in onda del secondo videoclip. Esempi di padding clip sono il logo del servizio, oppure, nel caso del servizio pilota TG Regionali, informazioni sulla regione a cui le news seguenti si riferiranno. La Modalità Sincrona è particolarmente utile per creare servizi di cui sia necessario conoscere gli orari, come il servizio TG Regionali

- Modalità Mista: è possibile specificare l'orario di alcuni videoclip (Modalità Sincrona) e inserire altri elementi senza un orario specifico (Modalità Sequenziale), in modo che siano trasmessi sequenzialmente, in modalità carousel, fino al seguente elemento in Modalità Sincrona. Questa modalità è utile per creare carousel tematici in diverse time slot (es. UGC dalle 22:00 alle 23:00 e Show dalle 23:00 alle 23:30).

Esiste un'ulteriore modalità di trasmissione, la Modalità Video Blog, che consente di mandare in streaming sequenzialmente un numero prefissato di videoclip, in modalità carousel, con la possibilità di inserimento dinamico di nuovi video clip. La lunghezza del carousel è determinata dal numero di videoclip selezionati. L'attuale imple-

Fig. 8 – Pagina Web per il Video Blog.

mentazione prevede la selezione dei videoclip più recenti: se un nuovo videoclip è stato inserito nel carousel, il più vecchio viene eliminato. Quando un nuovo videoclip è disponibile, perché appena selezionato tra quelli inviati dagli utenti, gli viene assegnata una priorità per la successiva schedulazione. Si è deciso di dare ai clip appena inseriti priorità massima: in questo modo verranno schedulati subito.

In questo modo, la Modalità Video Blog permette la trasmissione continua e l'aggiornamento dinamico del palinsesto. Una caratteristica saliente della Modalità Video Blog è la possibilità di associare un videoclip introduttivo ad ogni videoclip UGC. Questo videoclip introduttivo (Title Clip nel seguito) verrà trasmesso immediatamente prima del videoclip UGC a cui si riferisce, e conterrà, ad esempio, il logo del servizio, alcune informazioni sul videoclip UGC e una lista dei videclip che seguiranno. Il sistema prototipale realizzato permette la creazione automatica dei Title Clip, partendo dai metadati associati a ciascun clip UGC, cioè titolo, descrizione, categoria, durata e data di creazione; la lista di schedulazione prelevata dallo Schedulatore permette inoltre di inserire una breve lista dei clip seguenti.

Per tutti i servizi pilota realizzati con il sistema prototipale (per ora il servizio Regionale e il Video Blog), contemporaneamente alla creazione delle playlist per il broadcast, viene aggiornato anche un sito web mobile, accessibile via GPRS/UMTS/HSDPA che conterrà una selezione molto più ampia dei videclip, insieme agli stessi metadati. Si veda in figura 8 un esempio della presentazione Mobile Web del Video Blog.

Lo schedulatore può essere controllato in remoto utilizzando la sua interfaccia http. Da qui è possibile seguire l'andamento delle schedulazioni con un Monitor (figura 9, in cui l'elemento evidenziato è in streaming)

Si noti che nel Servizio Pilota è stata implementata una strategia di schedulazione molto semplice,

Fig. 9 – Interfaccia dello schedulatore: sequenza di schedulazione.

Titolo	Categoria	Durata	Data
Degrado dell'ex centrale del latte di Firenze	Informazione	0:36	21/03/2008
Farmacia Chiusa..	Amatoriale	1:47	19/03/2008
Milano, mercato di via Papiniano	Amatoriale	0:55	16/03/2008
Vittoria della Juve a Milano	Sport	1:23	06/03/2008
Parcheggi devastati	Amatoriale	0:48	28/02/2008
Sottopasso di Roma	Amatoriale	0:35	24/02/2008
Incendio in una palazzina a Torino	Informazione	1:25	04/02/2008

che consiste in un First In First Out (FIFO) per i clip già inseriti nel palinsesto, e in un Last In First Out (LIFO) per i clip nuovi, presumibilmente più interessanti per gli utenti. Si è quindi ipotizzato che il numero di clip inseriti quotidianamente sia molto più piccolo dei clip già in onda.

7. Scalabilità del sistema

Relativamente alla scalabilità del sistema, il Servizio Pilota più critico è sicuramente il Video Blog, in quanto il numero di videoclip ricevuti e la loro schedulazione sono potenzialmente imprevedibili. Il sistema è stato messo alla prova con reali UGC inseriti da utenti e, in generale, può funzionare senza problemi con diverse decine di videoclip al giorno. Un servizio molto popolare avrebbe probabilmente diverse centinaia di video blogger pronti a contribuire con i loro filmati, e per fronteggiare a situazioni simili il sistema è stato progettato per essere facilmente espanso. Prenderemo in esame nel seguito alcuni punti critici riguardanti la scalabilità.

In primo luogo, il sistema richiede un numero molto limitato di encoder. Le richieste vengono serializzate e viene utilizzato un solo encoder H.264 per servizio. Se il numero di video ricevuti dagli utenti aumentasse e fosse necessario mandare in onda i videoclip con i relativi Title Clip in breve tempo, due encoder sarebbero più che sufficienti per un singolo servizio.

In secondo luogo, lo Schedulatore non necessita di modifiche, se non nelle strategie di messa in onda, al crescere del numero dei videoclip: ogni servizio A/V disporrà di un proprio Schedulatore.

In terzo luogo, la parte Mobile Web del sistema è costituita da clip compressi (pochi MB generalmente): i moderni sistemi di storage permettono di archiviarne centinaia di migliaia.

8. Conclusioni

In questo articolo è stato descritto un innovativo sistema prototipale per la creazione automatica, la pubblicazione e la messa in onda di canali “tematici” specifici per l’impiego su piattaforme di TV mobile. Il sistema permette di recuperare ed elaborare automaticamente materiale audio-video d’archivio, oppure generato dagli utenti, e predisporne lo streaming secondo un palinsesto definibile attraverso diverse strategie di schedulazione. Vengono illustrati gli scenari applicativi che trovano applicazione come due canali nel pacchetto di servizi di TV Mobile diffuso in DVB-H nell’area di Torino: un canale di News Regionali (carousel) e un canale Video Blog basato su User Generated Contents. Oltre alla diffusione sulla rete DVB-H il sistema prototipale descritto consente anche la pubblicazione dei contenuti su un sito “Mobile Web”, realizzando così un primo passo nell’integrazione di diversi “nuovi media”.

Bibliografia

1. Sito DVB-H Global Mobile TV, ultimo accesso 1 Aprile 2008, www.dvb-h.org
2. G. Alberico, P. Casagranda, C. Migliardi: “DVB-H e Mobile TV: nuovi contenuti per nuove modalità di fruizione”, Elettronica e Telecomunicazioni, ERI, Aprile 2007
3. Dr. Shani Orgad: “This box was made for walking”, London School of Economics, Novembre 2006
4. “Transmission System for Handheld Terminals (DVB-H)”, ETSI EN 302 304, ETSI, Novembre 2004
5. “DVB-H Implementation Guidelines”, ETSI, ETSI TR 102 377, Novembre 2005

6. “DVB Specification for data broadcasting”, ETSI EN 301 192, ETSI, Novembre 2004

7. “IP Datacast over DVB-H: Electronic Service Guide (ESG)”, ETSI TS 102 471, ETSI, Novembre 2006

8. T. Paila, M. Luby, R. Lehtonen, V. Roca, R. Walsh: “FLUTE - File Delivery over Unidirectional Transport”, RFC 3926, IETF, Ottobre 2004

9. Tamura, H., Mori, S., Yamawaki, T.: Texture features corresponding to visual perception. IEEE Trans. on Systems, Man and Cybernetics, 8(6):460-473, 1978

Distorsioni nella stereoscopia tradizionale

ing. Mario **Muratori**
Rai
Centro Ricerche e
InnovazioneTecnologica
Torino

1. Introduzione

La filiera di produzione stereoscopica, anche se limitata all'essenziale, cioè al sistema di ripresa e al sistema di visualizzazione, non è in grado di visualizzare la scena ripresa senza distorsioni, a meno di non adottare configurazioni particolari.

Alcune distorsioni dipendono dalla configurazione geometrica della ripresa, altre dalle caratteristiche delle ottiche, altre ancora dipendono dal metodo di visualizzazione.

In [1] si analizzano geometricamente le configurazioni di ripresa e si introducono le relazioni che legano le coordinate dei punti della scena reale con quelle dei punti percepiti nello spazio visualizzato, che sono utilizzate in questo lavoro per dare giustificazione matematica ad alcune distorsioni.

Sommario
La filiera di produzione stereoscopica non è in grado di visualizzare la scena ripresa senza distorsioni, a meno di non adottare particolari configurazioni del sistema. Alcune distorsioni dipendono dalla configurazione geometrica della ripresa, altre dalle caratteristiche delle ottiche, altre ancora dipendono dal metodo di visualizzazione. In questo lavoro si presenta una panoramica delle principali distorsioni da cui è affetta la tecnica stereoscopica tradizionale[Nota 1]*, che vengono analizzate al fine di individuare eventuali misure correttive.*

Nota 1 - Per stereoscopia tradizionale si intende l'insieme di tecniche basate su due sensori che generano una coppia di segnali (coppia stereoscopica) e le relative tecniche di visualizzazione.

Si analizzano inoltre anche altri fenomeni, con l'obiettivo di presentare una panoramica delle principali distorsioni cui è affetta la tecnica stereoscopica tradizionale, analizzarle ed individuare eventuali misure correttive.

2. Relazione tra lo spazio percepito e lo spazio della scena

Per analizzare le distorsioni tipiche della stereoscopia, è necessario mettere in relazione lo spazio tridimensionale della scena ripresa, in cui sono definite le coordinate X_p, Y_p, Z_p, con lo spazio tridimensionale percepito dall'osservatore in fase di visualizzazione, nel quale sono definite le coordinate X_o, Y_o, Z_o.

Questo si ottiene combinando le relazioni che descrivono i rapporti tra lo spazio della scena e la sua proiezione sui sensori, le relazioni esistenti tra le coordinate dello schermo e le coordinate sui sensori, e le relazioni che legano le proiezioni dei punti sullo schermo con la posizione percepita dall'osservatore nello spazio virtuale generato dalla visione stereoscopica.

Si ottengono le seguenti relazioni, valide per la configurazione ad assi ottici convergenti:

1a)
$$X_o = \frac{Mf}{2} \frac{tg\left[arctg\left(\frac{2X_p + b}{2Z_p}\right) - \beta\right] + tg\left[arctg\left(\frac{2X_p - b}{2Z_p}\right) + \beta\right]}{1 + \frac{Mf}{e}\left\{tg\left[arctg\left(\frac{2X_p + b}{2Z_p}\right) - \beta\right] - tg\left[arctg\left(\frac{2X_p - b}{2Z_p}\right) + \beta\right]\right\}}$$

1b)
$$Y_o = 2MfY_p \frac{\frac{2Z_p \cos\beta + b \sin\beta}{\left[2Z_p \cos\beta + (2X_p + b)\sin\beta\right]\left[2Z_p \cos\beta - (2X_p - b)\sin\beta\right]}}{1 + \frac{Mf}{e}\left\{tg\left[arctg\left(\frac{2X_p + b}{2Z_p}\right) - \beta\right] - tg\left[arctg\left(\frac{2X_p - b}{2Z_p}\right) + \beta\right]\right\}}$$

1c)
$$Z_o = \frac{Z_{oS}}{1 + \frac{Mf}{e}\left\{tg\left[arctg\left(\frac{2X_p + b}{2Z_p}\right) - \beta\right] - tg\left[arctg\left(\frac{2X_p - b}{2Z_p}\right) + \beta\right]\right\}}$$

Ponendo β = 0 nelle relazioni precedenti si ricavano le relazioni valide per la configurazione ad assi ottici paralleli, che si esplicitano perché ampiamente utilizzate nel seguito:

2a)
$$X_o = \frac{MfX_p}{Z_p + \frac{Mfb}{e}}$$

2b)
$$Y_o = \frac{MfY_p}{Z_p + \frac{Mfb}{e}}$$

2c)
$$Z_o = \frac{Z_{oS} Z_p}{Z_p + \frac{Mfb}{e}}$$

Si noti la similitudine formale tra le relazioni 2a) e 2b). Interessante è anche l'uguaglianza del polinomio al denominatore nelle tre le relazioni, poiché esso è un elemento caratterizzante l'andamento delle funzioni rispetto alla coordinata Z_p.

3. Dominio di validità delle relazioni

Si considerino le relazioni che legano le coordinate dei punti immagine proiettate sullo schermo, $x_{Sd,s}$ e $y_{Sd,s}$, con quelle dei punti costituenti la scena percepita, X_o, Y_o, Z_o:

3a) $$X_o = -\frac{(x_{Sd} + x_{Ss})}{2}\frac{e}{e-\delta_{So}} = -x_{Smed} * R$$

3b) $$Y_o = \frac{(y_{Sd} + y_{Ss})}{2}\frac{e}{e-\delta_{So}} = y_{Smed} * R$$

3c) $$Z_o = Z_{oS}\frac{e}{e-\delta_{So}} = Z_{oS} * R$$

Dove si è introdotto il rapporto R:

4) $$R = \frac{e}{e-\delta_{So}}$$

Come funzione della disparità, *R* è rappresentato da un ramo di iperbole traslata lungo le ascisse del valore *e*, come illustrato in figura 1.

Il rapporto *R* deve essere positivo per non cambiare il segno alle coordinate percepite; tenendo conto che la distanza interoculare è sempre positiva, le considerazioni precedenti implicano le condizioni seguenti:

5a) $\delta_{So} \leq e$ ovvero: 5b) $\delta_o \leq e/M$

che hanno un significato fisico evidente: il valore della disparità orizzontale dei punti proiettati sullo schermo non può superare quello della distanza interoculare, altrimenti si richiederebbe la divergenza degli assi ottici oculari in fase di visione. La condizione di uguaglianza corrisponde alla situazione limite per cui il punto visualizzato viene percepito a distanza infinita, e che è rappresentata dal polo in $\delta_{So} = e$.

In alcuni casi piuttosto interessanti discussi nel seguito, si traslano le immagini proiettate sullo schermo della distanza interoculare. In formule si tratta di utilizzare delle ascisse (definite sullo schermo), indicate con $x'_{Sd,s}$ modificate come segue:

6a) $x'_{Sd} = x_{Sd} + e/2$ 6b) $x'_{Ss} = x_{Ss} - e/2$

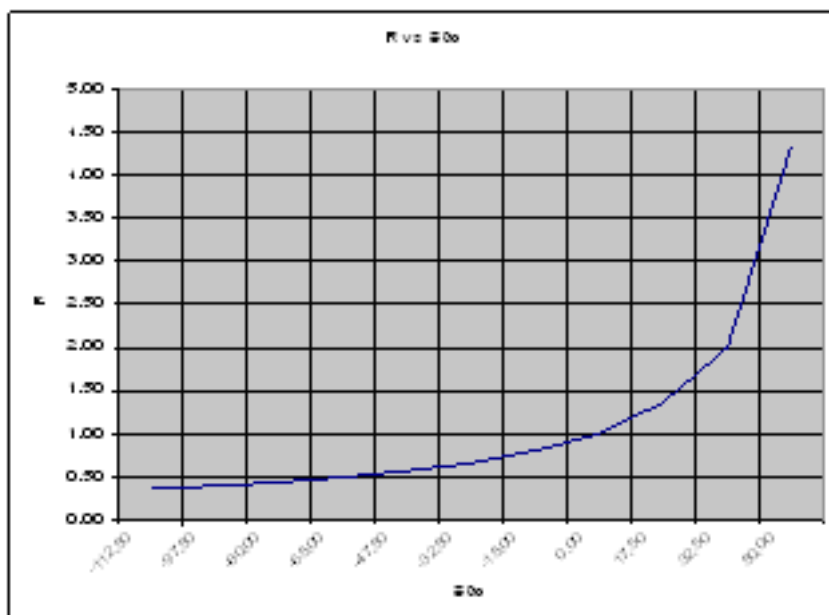



Fig. 1 – Andamento di $R(\delta_{So})$ con e = 65 mm

Le relazioni 3a), 3b), 3c) si modificano di conseguenza come segue:

7a) $$X_o = -\frac{(x_{Sd} + x_{Ss})}{2}\frac{e}{e-\delta_{So}} = -x_{Smed} * R'$$

7b) $$Y_o = \frac{(y_{Sd} + y_{Ss})}{2}\frac{e}{e-\delta_{So}} = y_{Smed} * R'$$

7c) $$Z_o = Z_{oS}\frac{e}{e-\delta_{So}} = Z_{oS} * R'$$

Con *R'* dato dalla relazione seguente:

8) $$R' = -\frac{e}{\delta_{So}}$$

In funzione della disparità δ_{So}, *R'* è rappresentato da un ramo di iperbole definito nel dominio $\delta_{So} < 0$, e che presenta un polo in $\delta_{So} = 0$.

Anche queste posizioni hanno un significato fisico chiaro: a causa della traslazione imposta alle immagini proiettate, il valore nullo della disparità orizzontale corrisponde alla percezione "all'infinito" del punto considerato. Valori positivi della disparità implicano la divergenza degli assi ottici oculari.

Quando si superano i limiti suindicati, alcuni punti della scena reale, in linea di massima ubicati ai margini laterali della scena ripresa, sono visualizzati in maniera non corretta e generano diplopia.

In pratica, tuttavia, ciò potrebbe non essere un problema rilevante, in quanto l'ampiezza di campo degli obiettivi tende a escludere la parte di scena che potenzialmente potrebbe non essere correttamente visualizzata, ed inoltre nell'apparato di visione dell'osservatore si attiva la funzionalità di cancellazione dei segnali ottici che danno diplopia.

4. Analisi delle distorsioni generate in fase di ripresa

4.1 Non linearità nella riproduzione delle posizioni nello spazio

Le relazioni riportate in precedenza mostrano che il valore delle coordinate X_o, Y_o, Z_o non è direttamente proporzionale a quello della coordinata Z_p, come invece ci si aspetterebbe se valessero le leggi della prospettiva. Infatti le loro derivate rispetto a Z_p risultano proporzionali al rapporto:

9) $$R_d = \frac{1}{(Z_p + \frac{Mfb}{e})^2}$$

che non ha un valore costante in funzione di Z_p.

Nel caso di assi ottici convergenti la derivazione matematica è più complessa, ma porta alla stessa conclusione generale sulla mancanza di proporzionalità diretta.

Non devono stupire quindi gli andamenti illustrati nelle figure 2. Questi inoltre evidenziano che nella configurazione ad assi convergenti non è remota la possibilità di ottenere divergenza degli assi ottici oculari in fase di visualizzazione, testimoniata dalla presenza del polo e delle associate irregolarità dei grafici.

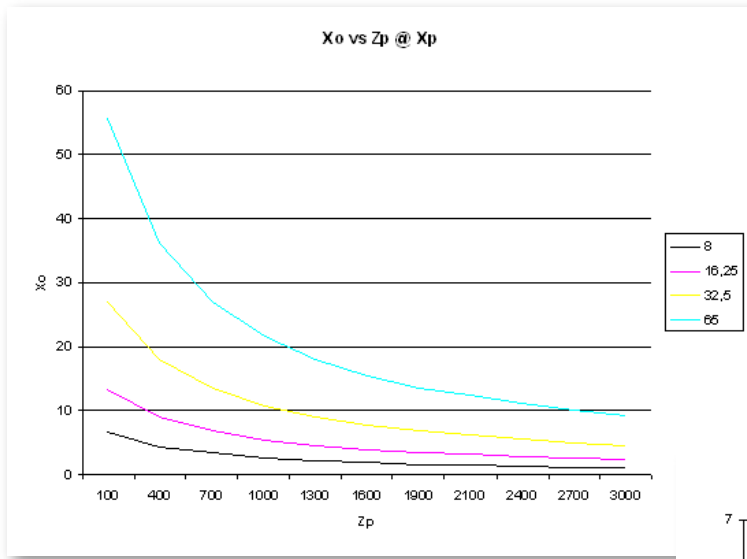
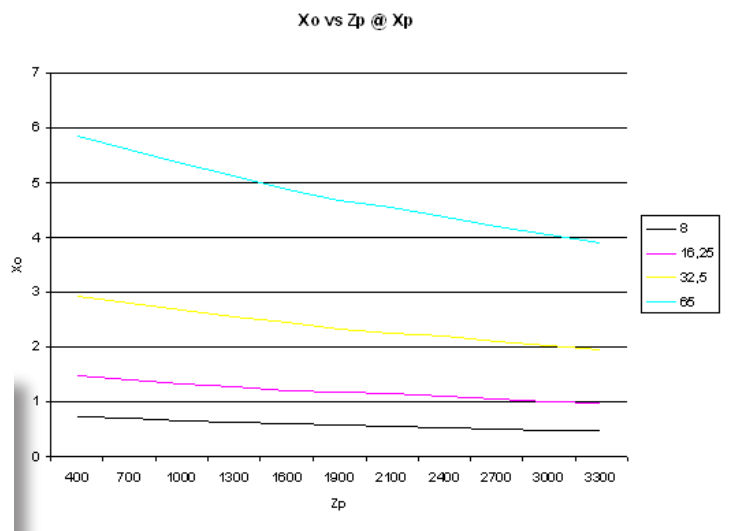
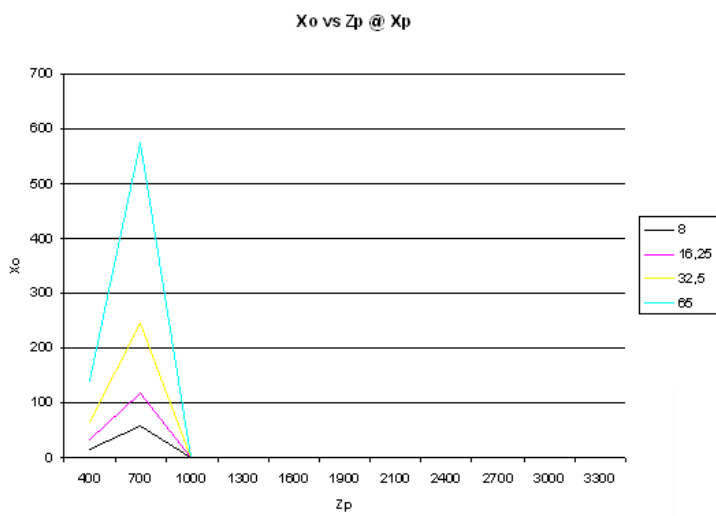
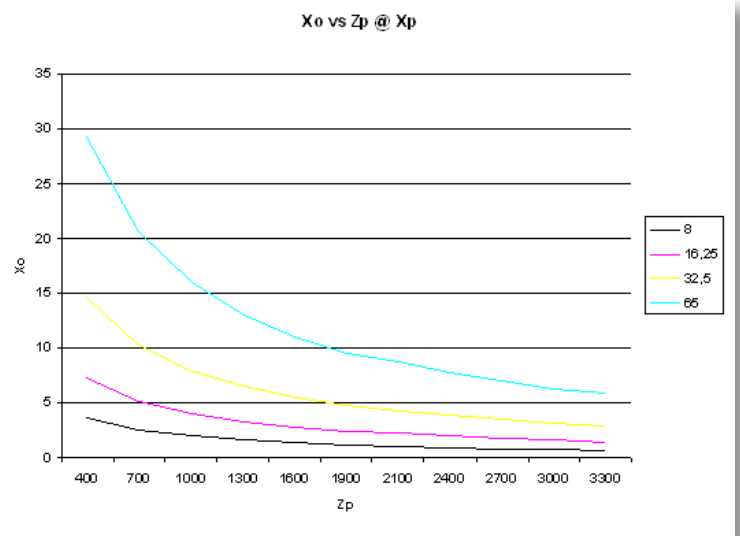



Fig. 2a - Andamento di X_o in funzione di Z_p, per alcuni valori di X_p

a) b = 65 mm, f = 6.5 mm, β =2°, M = 50

b) b = 650 mm, f = 6.5 mm, β =2°, M = 50

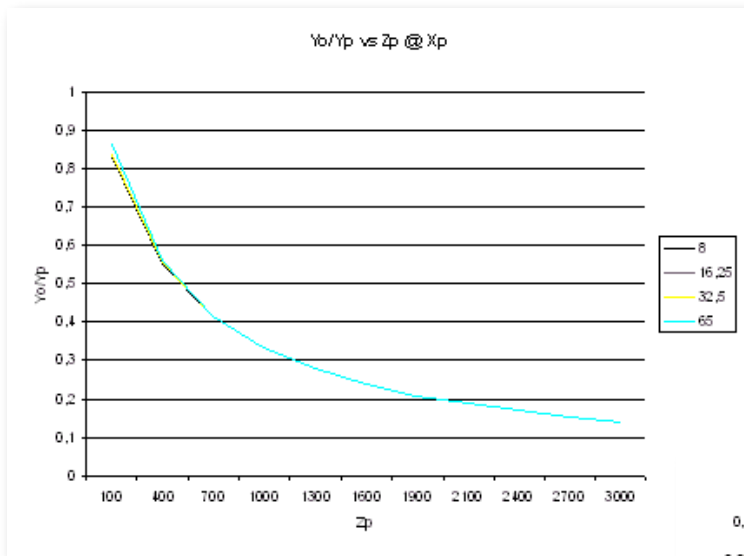
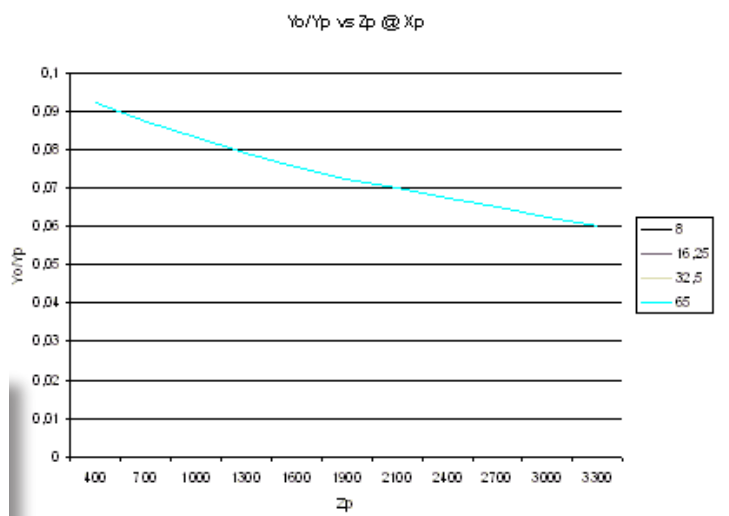
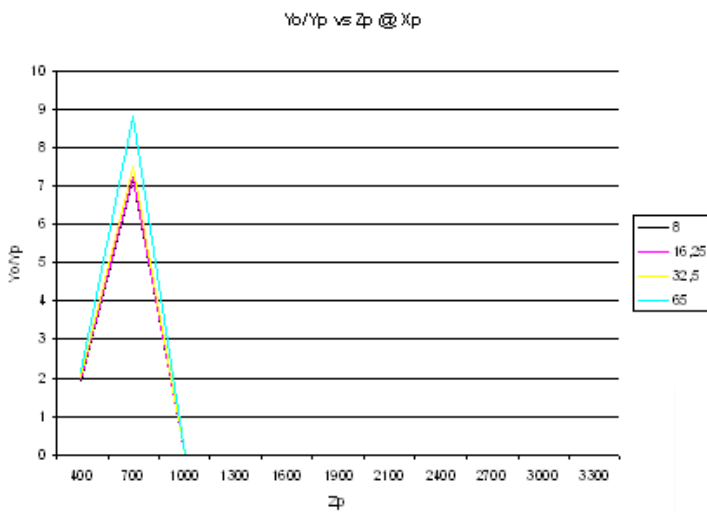
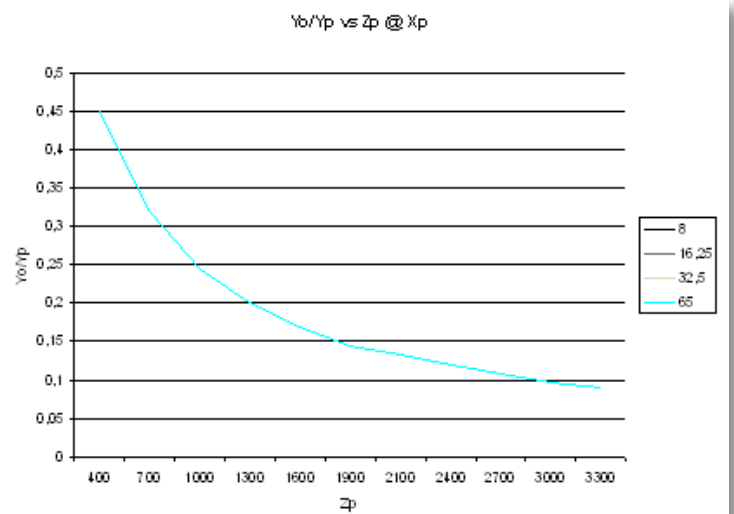
c) b = 65 mm, f = 6.5 mm, β =8°, M = 50

d) b = 65 mm, f = 6.5 mm, β =0°, M = 50

Fig. 2b - Andamento di Y_o in funzione di Z_p, per alcuni valori di X_p, con Y_p=100.

a) b = 65 mm, f = 6.5 mm, β =2°, M = 50

b) b = 650 mm, f = 6.5 mm, β =2°, M = 50

c) b = 65 mm, f = 6.5 mm, β =8°, M = 50

d) b = 65 mm, f = 6.5 mm, β =0°, M = 50

4.2 Non linearità nella riproduzione della profondità

Le figure 3a e 3b illustrano l'andamento della coordinata Z_o del punto percepito in funzione della coordinata Z_p del corrispondente punto della scena reale.

Per la precisione, nei grafici si riporta l'andamento del rapporto tra distanza percepita Z_o e distanza dello schermo Z_{oS}, in funzione della coordinata Z_p; inoltre, si è posto per semplicità, ma senza perdere in significato, $X_p = Y_p = 0$.

La relazione tra le due grandezze è non lineare, come emerge anche dall'analisi della derivata rispetto a Z_p della relazione 2c).

In particolare, per ridotte lunghezze focali, la distanza relativa tra oggetti vicini viene percepita dilatata rispetto a quella reale; il contrario si verifica per gli oggetti lontani, la cui distanza relativa viene percepita compressa rispetto alla realtà.

A causa di questa distorsione, la valutazione della distanza e di altre grandezze derivate, quali la velocità, tramite sistema di ripresa stereoscopico può risultare erronea[Nota 2].

Nel caso della configurazione ad assi ottici paralleli si ottiene che per Z_p tendente all'infinito, Z_o tende alla distanza tra osservatore e schermo, come prevedibile, benché percettivamente errato.

Invece, analizzando il grafico di figura 3b, relativo alla configurazione ad assi ottici convergenti, emerge che esiste un valore di Z_p per cui la distanza percepita diventa uguale

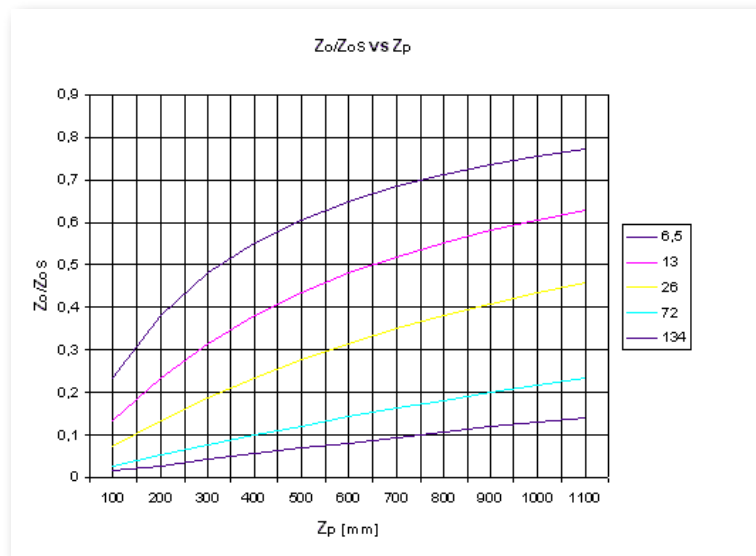



Fig. 3a - Relazione tra la distanza percepita in funzione di quella reale - configurazione ad assi paralleli. Parametro lunghezza focale f; e = 65 mm, b = 65 mm, M = 50

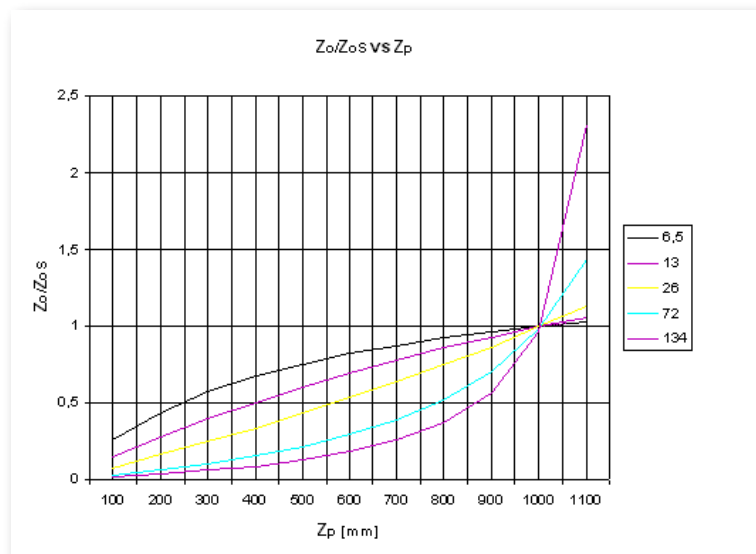



Fig. 3b - Relazione tra la distanza percepita in funzione di quella reale - configurazione ad assi convergenti. Parametro lunghezza focale f; e = 65 mm, b = 65 mm, M = 50, β = 1.85°.

Nota 2 - Questo effetto genera difficoltà nel controllo remoto di mezzi semoventi qualora si utilizzi un sistema stereoscopico per la ripresa e la visualizzazione dell'ambiente dove opera il mezzo.

alla distanza dello schermo dall'osservatore (Z_o/Z_{oS} = 1); in questo caso, il punto (0, 0, Z_p) coincide col punto di convergenza degli assi ottici delle camere[Nota 3].

Non rappresentato in figura, esiste un altro valore di Z_p caratteristico - nella configurazione ad assi convergenti - che corrisponde al punto le cui immagini sono proiettate sullo schermo con disparità pari alla distanza interoculare. In questo caso la distanza percepita Z_o tende ad infinito sebbene i corrispondenti punti della scena reale abbiano in genere valore di Z_p finito, confermando ulteriormente la non linearità della relazione tra le due grandezze. I punti con valori di Z_p ancora maggiore sono proiettati sullo schermo con una disparità tale da richiedere, per essere osservati contemporaneamente, la divergenza degli assi ottici oculari, impossibile ad ottenersi per definizione. In questo caso si verifica diplopia, cui normalmente consegue la cancellazione automatica ed inconscia di una delle due componenti stereoscopiche[Nota 4].

La percezione non lineare della distanza fa sì che la valutazione della stessa basata sul confronto delle dimensioni possa essere in disaccordo con quella basata sulla disparità orizzontale, generando un senso di confusione[Nota 5]. Per lo stesso motivo si può incorrere in un'erronea valutazione delle dimensioni dell'oggetto visualizzato, oltre che della sua distanza. Inoltre, a questo effetto si attribuisce anche la causa della distorsione chiamata *puppet-theatre effect* discussa nel seguito.

4.3. Non linearità nella riproduzione delle dimensioni traversali e riproduzione in scala

Le dimensioni trasversali sono valutabili calcolando la differenza tra le coordinate dei punti posti alle estremità degli oggetti considerati: la larghezza è data dalla differenza tra ascisse, l'altezza dalla differenza tra ordinate, la profondità dalla differenza tra coordinate Z.

In particolare, nella configurazione ad assi paralleli, si ottiene:

10a) $$\frac{\Delta X_o}{\Delta X_p} = \frac{Mf}{Z_p + \frac{Mfb}{e}}$$

10b) $$\frac{\Delta Y_o}{\Delta Y_p} = \frac{Mf}{Z_p + \frac{Mfb}{e}}$$

10c) $$\frac{\Delta Z_o}{\Delta Z_p} \cong \frac{Mfb}{e} \frac{Z_{oS}}{(Z_p + \frac{Mfb}{e})^2},$$

con $\Delta Z_p << Z_p$

Come si nota, la non linearità discussa nel paragrafo precedente si ripercuote anche nelle differenze tra coordinate, sicché si può affermare che anche la riproduzione delle dimensioni non

Nota 3 - Si ricorda che, nella derivazione dei grafici, si è posto X_p = Y_p = 0 per semplicità di trattazione.
Nota 4 - Peraltro questa è una situazione normale nel processo di visione umana, in cui gli assi ottici oculari convergono verso il punto di attenzione.
Nota 5 - Si ricorda che il confronto tra le dimensioni di oggetti noti è una delle tecniche per la valutazione della distanza messe in atto dal processo della visione umana, così come l'analisi delle disparità orizzontali e altri meccanismi che operano contemporaneamente. Tuttavia se le valutazioni ottenute non sono congruenti si genera un senso di sconforto e confusione. Alcuni artisti e ricercatori utilizzano tali incongruità per ottenere gli effetti artistici desiderati o le sensazioni fisiologiche volute (illusioni ottiche).

sia lineare, ma segua un andamento dello stesso tipo di quello illustrato per i valori delle singole coordinate, generando delle distorsioni di forma negli oggetti visualizzati.

Nel caso di configurazione ad assi ottici paralleli, applicando le posizioni rappresentate dalle relazioni 6a) e 6b) sopra riportate, nonchè la seguente:

11) $M f = Z_{oS}$

le relazioni che legano le coordinate della scena reale con quelle della scena percepita diventano le seguenti:

12a) $X_0 = eX_p/b$

12b) $Y_0 = eY_p/b$

12c) $Z_0 = eZ_p/b$

le quali indicano come, in questo caso, lo spazio percepito dall'osservatore in fase di visualizzazione, coordinate (X_o, Y_o, Z_o), risulti una riproduzione in scala, senza distorsioni di forma, dello spazio tridimensionale della scena ripresa, coordinate (X_p, Y_p, Z_p).

4.4 Effetto teatrino (puppet-theatre effect)

Non è ancora stata trovata una definizione condivisa per l'effetto teatrino, noto col termine inglese *puppet-theatre effect*; in letteratura si trovano lavori in cui ne viene data una definizione per poterne quantificare l'entità (per esempio [2]).

In ogni caso, si tratta di una distorsione dimensionale per cui gli oggetti visualizzati in un sistema stereoscopico appaiono innaturalmente piccoli; di conseguenza, la visualizzazione perde verosimiglianza e appare artefatta.

Per semplicità di trattazione si considera il caso di assi ottici paralleli ($\beta = 0°$).

Dall'analisi degli andamenti riportati nelle figure 2 e 3, si desume che, nei campi vicini (corta-media distanza), le dimensioni degli oggetti visualizzati vengono percepite più piccole di quello che dovrebbero essere se valessero le leggi lineari della prospettiva, mentre le distanze sono percepite dilatate.

Al contrario, nei campi lontani (distanze medio-lunghe) l'andamento delle dimensioni è asintotico e le distanze sono percepite compresse.

Ciò porta a vedere la scena come se fosse, appunto, in un teatro, dove gli oggetti posti a distanza ravvicinata sono rappresentati sul palco e vengono percepiti relativamente piccoli (perché confrontati con la scenografia del palcoscenico) e ben distanziati tra loro, in contrasto con quelli posti a maggiore distanza che risultano invece schiacciati sullo sfondo (nella pratica teatrale spesso dipinti su di esso).

Questa analisi sembra coincidere con quanto espresso in [2].

Si può affermare quindi che l'effetto teatrino dipenda dalla riproduzione non lineare dello spazio; solo con l'adozione delle posizioni per la riproduzione in scala si ottiene un andamento lineare con la distanza, e quindi l'assenza dell'effetto. Nel caso della configurazione di ripresa ad assi ottici convergenti il verificarsi dell'effetto teatrino è molto probabile, dato che l'andamento dei grafici delle figure 2 e 3 non è sostanzialmente dissimile da quello ottenuto per la configurazione ad assi ottici paralleli, ma la presenza di irregolarità e di andamenti peculiari (per esempio per quanto riguarda la distanza percepita) indica che la situazione è meno predicibile e dipendente dal valore dei parametri di ripresa adottati.

4.5 Effetto cartolina (cardboard effect)

Questo effetto si verifica quando gli osservatori percepiscono gli oggetti non come solidi tridimensionali, ma come figure piatte, ancorché correttamente posizionate nel senso della profondità. Risulta fastidioso in quanto toglie verosimiglianza allo spettacolo visualizzato.

Anche questo effetto, come il precedente, non ha ancora trovato una definizione condivisa. In [2] se ne dà una definizione in termini di rapporto E tra il fattore di ingrandimento longitudinale Il e quello trasversale I_t, qui riportati nel caso di assi ottici paralleli:

13a) $$I_t = \frac{\Delta X_o}{\Delta X_p} = \frac{\Delta Y_o}{\Delta Y_p} = \frac{Mf}{Z_p + \frac{Mfb}{e}}$$

13b) $$I_l = \frac{\Delta Z_o}{\Delta Z_p} = \frac{Z_{oS}}{Z_p + \frac{Mfb}{e}}$$

13c) $$E = \frac{I_l}{I_t} = \frac{Z_{oS}}{Mf}$$

Se E è pari all'unità non si verifica alcuna distorsione dimensionale e gli oggetti vengono percepiti nella loro tridimensionalità, in accordo con la posizione 11) necessaria per ottenere una riproduzione in scala.

Invece se E risulta minore dell'unità significa che la riproduzione è compressa nella direzione delle profondità e quindi gli oggetti appaiono schiacciati e si verifica l'effetto *cardboard*. Tanto diverso dall'unità è il valore di E, tanto maggiore risulta la distorsione.

4.6 Curvatura del piano della profondità

Nella configurazione ad assi ottici convergenti si verifica la curvatura dei piani di profondità, che, invece, rimangono paralleli all'asse X_o nella configurazione ad assi ottici paralleli. L'effetto è illustrato nelle figure 4a e 4b.

A causa di questa distorsione, gli oggetti posti ai lati dell'immagine vengono percepiti più distanti di quelli che si trovano al centro dell'immagine, anche se nella scena reale si trovano con lo stesso valore della coordinata Z_p.

Pertanto questa distorsione induce una errata valutazione della distanza relativa tra gli oggetti costituenti la scena visualizzata; inoltre può dare sensazioni fastidiose - sembra che la scena si muova - quando l'apparato di ripresa effettua un panning[Nota 6].

Nota 6 - Il panning è la rotazione attorno all'asse verticale dell'apparato di ripresa.

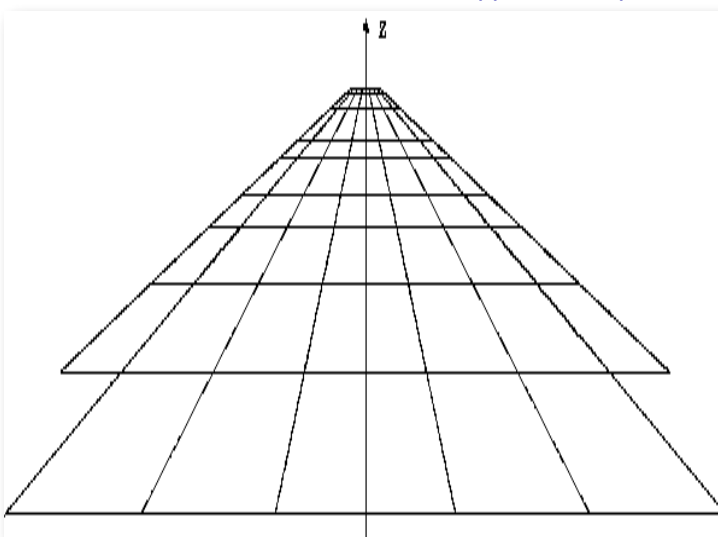



Fig. 4a - Proiezione nello spazio percepito in fase di visualizzazione di un reticolo ortogonale uniforme definito nello spazio della scena – configurazione ad assi ottici paralleli (vista dall'alto)

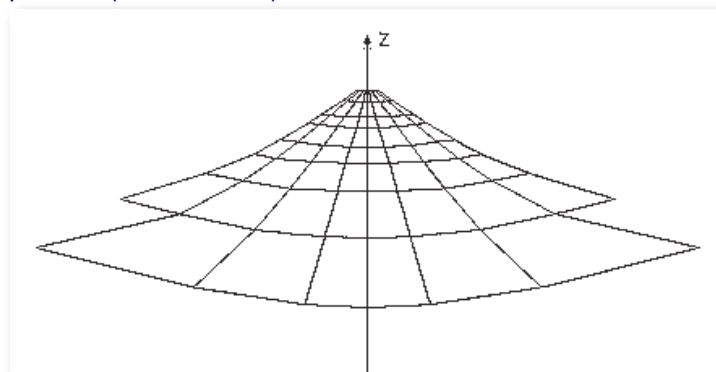



Fig. 4b - Proiezione nello spazio percepito in fase di visualizzazione di un retico-lo ortogonale uniforme definito nello spazio della scena – configurazione ad assi ottici convergenti (vista dall'alto).

Questa distorsione è legata alla distorsione trapezoidale che si verifica ruotando i sensori rispetto alla scena ed è quindi inevitabile se si utilizza la configurazione ad assi ottici convergenti.

4.7 Distorsione trapezoidale (key-stone distortion)

La distorsione trapezoidale è tipica della configurazione di ripresa ad assi convergenti - quella ad assi ottici paralleli infatti non introduce questo tipo di distorsione - ed è dovuta al fatto che, in genere, le immagini del punto della scena considerato, proiettate sui sensori, si trovano su piani differenti.

Fig. 5 - Andamento della disparità verticale percepita.

Tale distorsione genera una disparità verticale, β_{ov}, osservabile in fase di visualizzazione, il cui valore è valutabile tramite la relazione seguente.

13)
$$\delta_{ov} = \frac{-2fY_pX_p\sin\beta}{\left[Z_p\cos\beta+\left(X_p+\frac{b}{2}\right)\sin\beta\right]\left[Z_p\cos\beta-\left(X_p-\frac{b}{2}\right)\sin\beta\right]} *$$
$$* \frac{1}{\left\{1+\frac{Mf}{e}\left[tg\left(arctg\left(\frac{2X_p+b}{2Z_p}\right)-\beta\right)-tg\left(arctg\left(\frac{2X_p-b}{2Z_p}\right)+\beta\right)\right]\right\}}$$

La figura 5 riporta l'andamento di δ_{ov}.

La disparità:

- è massima agli angoli dello schermo, essendo proporzionale a Y_p e X_p
- aumenta all'aumentare dell'angolo di convergenza, β
- aumenta all'aumentare della lunghezza focale, *f*
- è nulla per i punti giacenti sul piano *XZ* (Y_p = 0), la cui immagine giace sull'asse x dei sensori
- è nulla per i punti giacenti sul piano *YZ* (X_p = 0), grazie alla simmetria della configurazione geometrica; si noti che tali punti vengono visualizzati sul piano Y_oZ_o

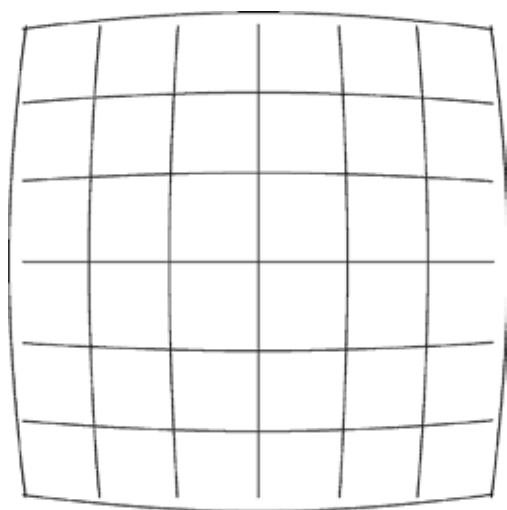
La distorsione trapezoidale introduce anche una disparità orizzontale spuria che si somma

a quella - desiderata - dovuta alla parallasse. Tale disparità è la causa della curvatura dei piani di profondità, di cui si è scritto nel paragrafo precedente.

4.8 Distorsioni dovute agli obiettivi

La distorsione radiale delle lenti, normalmente conosciuta come distorsione *a botte* o *a cuscino* a seconda di come si presenta, è un'altra sorgente di distorsione dell'immagine e di disparità verticale.

La distorsione radiale è causata dall'adozione di lenti sferiche, per esempio come elementi di un sistema ottico complesso, che presentano lunghezze focali diverse a seconda della distanza dall'asse ottico[Nota 7]. Se la lunghezza focale diminuisce all'aumentare della distanza dall'asse si verifica la distorsione a botte (figura 6a), al contrario, se la lunghezza focale aumenta la distorsione risultante è del tipo a cuscino (figura 6b).

Siccome queste distorsioni fanno sì che un reticolo ortogonale sia visualizzato deformato, punti omologhi aventi una disparità orizzontale di valore non nullo sono percepiti aventi ordinate diverse. In altre parole mostrano una disparità verticale indotta dalla distorsione delle lenti.

5. Analisi delle distorsioni generate in fase di visualizzazione

5.1 Distorsione per deformazione di taglio (shear distorsion)

Uno svantaggio dei visualizzatori stereoscopici binoculari è che la posizione dell'immagine percepita dipende da quella che ha l'osservatore rispetto allo schermo di visualizzazione[Nota 8].

La distanza percepita dipende linearmente dalla distanza tra osservatore e schermo, secondo la relazione 1c) e 2c). Allontanandosi dallo schermo, un punto percepito tra lo schermo e l'osservatore sembra spostarsi verso l'osservatore, mentre un punto percepito oltre lo schermo sembra allontanarsi ulteriormente, e viceversa.

La posizione del punto percepito, in termini di ascisse e ordinate, dipende dalla posizione che i punti proiettati sullo schermo hanno rispetto all'osservatore.

Nota 7 - Si veda anche: http://en.wikipedia.org/wiki/Image_distortion.
Nota 8 - Questo fenomeno non si verifica con i monitor autostereoscopici.

Fig. 6a - Distorsione *a botte*.

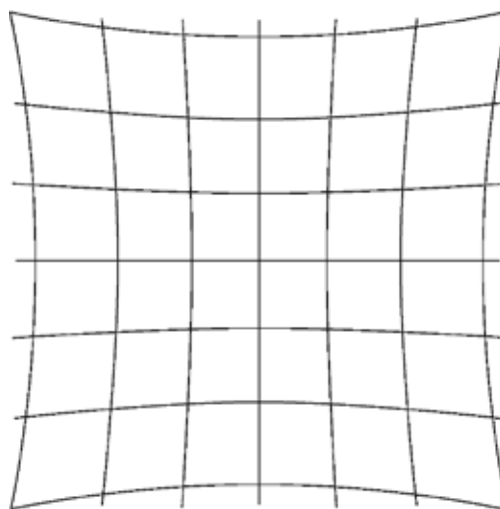



Fig. 6b - Distorsione *a cuscino*

Infatti, introducendo i termini x'_S e y'_S, che tengono conto della traslazione del centro dello schermo rispetto all'asse Z_o, le coordinate del punto proiettato diventano le seguenti:

15a) $x'_{Sd} = x_{Sd} + x'_S$; 15b) $x'_{Ss} = x_{Ss} + x'_S$

16a) $y'_{Sd} = y_{Sd} + y'_S$; 16b) $Y'_{Ss} = y_{Ss} + y'_S$

Inserendo nelle relazioni 5b) e 5c) di [1], al posto delle ascisse e ordinate del punto proiettato, il valore dato dalle relazioni sopra riportate, si ottiene:

16a) $$X'_o = X_o - x'_S \frac{e}{e - \delta_{So}}$$

16b) $$Y'_o = Y_o + y'_S \frac{e}{e - \delta_{So}}$$

Con le posizioni assunte in precedenza, valori positivi per x_S indicano una traslazione verso destra dello schermo rispetto all'osservatore; valori positivi per y_S ne indicano una traslazione verso l'alto. Coerentemente, l'osservatore percepirà il punto in oggetto più a destra ($X'_o < X_o$) e più in alto ($Y'_o > Y_o$).

Si noti che lo spostamento dello schermo a destra e in alto corrisponde allo spostamento dell'osservatore in senso opposto, ossia a sinistra e in basso rispetto all'asse Z_o.

Quindi, il dissamento dell'osservatore rispetto all'asse dello schermo implica un'errore nella percezione della posizione dei punti visualizzati; in particolare, l'eventuale movimento dell'osservatore implica la percezione di uno spostamento dei punti osservati che in realtà non si verifica.

5.2 Variazione delle dimensioni di oggetti che si muovono lungo l'asse delle profodità

Un effetto che si verifica nella visione stereoscopica di oggetti relativamente vicini che si muovono nella direzione dell'asse delle profondità, Z_p, consiste nel fatto che le dimensioni percepite di tali oggetti aumentano all'allontanarsi degli oggetti e, viceversa, diminuiscono al loro avvicinarsi, in contrasto con l'esperienza reale. La figura 7 illustra schematicamente la situazione.

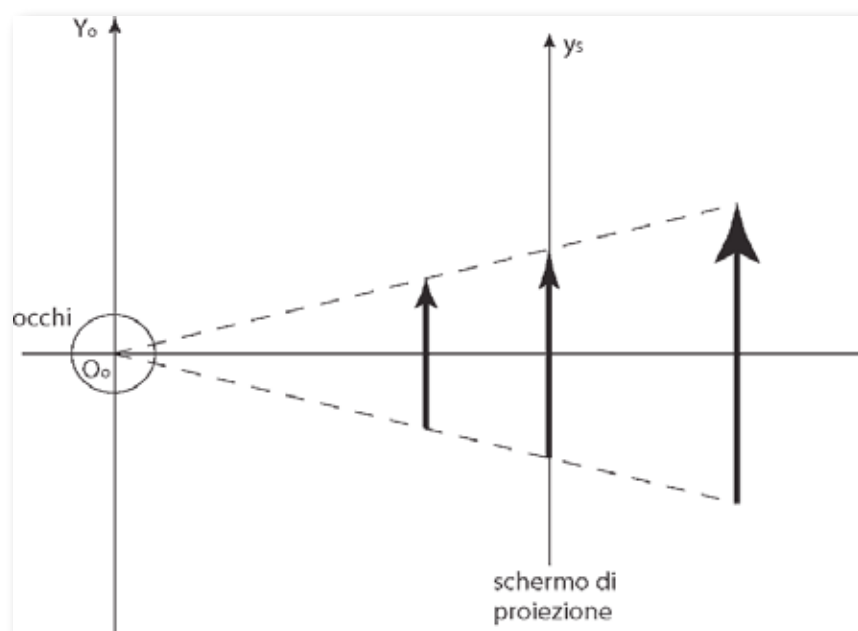



Fig. 7 - Illustrazione schematica dell'effetto di variazione delle dimensioni percepite di oggetti che traslano nella direzione delle profondità.

In caso di configurazione di ripresa ad assi paralleli, l'altezza e la larghezza percepite mostrano andamenti simultanei e di ugual valore; si può quindi, per semplicità, limitare l'analisi alla sola altezza senza perdere in generalità. Inoltre, le considerazioni ottenute per la configurazione ad assi ottici paralleli sono di principio valide anche per la configurazione ad assi ottici convergenti.

Le linee verticali illustrate in figura 7 presentano la stessa altezza sullo schermo di proiezione, ma a causa della loro disparità orizzontale sono percepite posizionate in diverse ubicazioni, cui corrisponde la percezione di altezze (in genere, di dimensioni) differenti. L'effetto è particolarmente visibile quando gli oggetti ripresi traslano, allontanandosi o avvicinandosi, rispetto alle camere.

Nel caso in cui l'oggetto ripreso si avvicini alla camera, le sue dimensioni aumentano, compensando parzialmente il rimpicciolimento percepito; viceversa qualora l'oggetto ripreso si allontani.

L'effetto complessivo è la combinazione di questi due effetti contrastanti.

Calcolando il rapporto tra le altezze percepite delle frecce poste a distanza differente, si ottiene:

18) $$\frac{Y_{o2}}{Y_{o1}} = \frac{Z_{p1} + \frac{Mfb}{e}}{Z_{p2} + \frac{Mfb}{e}}$$

che differisce dal semplice rapporto di proporzionalità diretta che ci si aspetterebbe valessero le leggi lineari della prospettiva. In altre parole, l'altezza relativa percepita tra oggetti posti a differente distanza dal sistema di ripresa non varia come nella realtà, a meno di non applicare le posizioni per la riproduzione in scala. In questo caso infatti si ottiene:

19) $$\frac{Y_{o2}}{Y_{o1}} = \frac{Z_{p1}}{Z_{p2}}$$

e quindi la variazione delle altezze percepite risulta simile a quella della realtà.

Considerazioni analoghe valgono anche per la dimensione trasversale.

5.3 Accomodazione e vergenza

Nel mondo reale, la convergenza oculare e l'accomodazione sono azioni strettamente legate fra loro, mentre nella visione stereoscopica richiedono interventi separati da parte dell'apparato di visione. Infatti, gli occhi mettono a fuoco la superficie dello schermo a prescindere dalla convergenza stimolata dalle disparità orizzontali proiettate.

Può quindi succedere che in corrispondenza a disparità orizzontali di valore elevato si perda la focalizzazione dell'immagine, oltre alla possibilità di non riuscire a fondere le due immagini.

In [4] si riportano i risultati di una analisi specifica. Riassumendo, si è trovato che capacità di visione stereoscopica differiscono grandemente da soggetto a soggetto. Alcuni tollerano disparità orizzontali di valore limitato, altri, al contrario vedono bene anche con valori molto elevati di disparità; alcuni riescono addirittura a divergere gli assi oculari. Vi sono alcuni soggetti che riescono a visualizzare meglio i soggetti percepiti "al di là" del monitor, altri viceversa visualizzano meglio quelli percepiti "al di qua" dello schermo.

In ogni caso sembra che l'abitudine alla visione stereoscopica ne aiuti la visualizzazione.

I risultati della ricerca suggeriscono di limitare la differenza tra le distanze degli oggetti ripresi – e conseguentemente visualizzati – per permettere a quanti più osservatori possibile di non incorrere in difficoltà di visione. Se la cosa non risulta possibile si deve quanto meno fare in modo che la parte della scena ripresa di maggiore interesse sia percepita nei pressi della superficie dello schermo, per esempio regolando opportunamente la distanza di convergenza degli assi ottici delle telecamere.

5.4 Parallasse verticale

Il parallasse verticale in fase di visualizzazione è causato dalle disparità verticali introdotte da alcune distorsioni.

In [4] si riporta anche il risultato di un esperimento volto a determinarne il massimo valore ammissibile, che risulta molto ridotto, dell'ordine di mezzo grado .

Sembrerebbe quindi buona norma minimizzare il parallasse verticale, operando su tutte le possibili fonti di distorsione, ad esempio la configurazione di ripresa, la lunghezza focale e gli obiettivi.

6. Considerazioni

Le cause principali di distorsione risiedono nella non linearità del legame tra spazio ripreso e spazio percepito e nell'angolazione dei sensori.

Schematizzando, si può affermare che la prima provochi:

- la riproduzione non lineare con la distanza delle dimensioni e della posizione degli oggetti ripresi, con conseguenti possibili errori di valutazione di questi parametri e incongruenze sensoriali;
- il *puppet-theatre effect* ed il *cardboard effect*, che diminuiscono la verosomiglianza della scena visualizzata;
- la percezione scorretta della dimensione degli oggetti in movimento lungo l'asse *Z*, che provoca incongruenze sensoriali.

La seconda provoca, in particolare:

- la generazione di disparità verticale spuria, che genera disturbo ed affaticamento visivo;
- la curvatura del piano delle profondità, che causa errori nella percezione della posizione degli oggetti e distorsioni di forma.

Inoltre si possono verificare altre distorsioni dovute alle ottiche e alle condizioni di visualizzazione.

In linea di principio, le distorsioni derivanti dalle non linearità potrebbero essere minimizzate considerando domini limitati, tali da poter linearizzare localmente le relazioni considerate. Questa operazione implica la riduzione dell'estensione dello spazio reale ripreso, cosa che coinvolge le scelte scenografiche e di fotografia.

Le distorsioni derivanti dall'inclinazione dei sensori sono eliminate adottando la configurazione di ripresa ad assi ottici paralleli, la quale peraltro è condizione necessaria ma non sufficiente ad assicurare una riproduzione in scala della scena ripresa. In ogni caso, limitare l'angolo di incidenza degli assi ottici dei sensori sembra essere una buona politica.

7. Conclusioni

La visualizzazione stereoscopica è affetta da diverse distorsioni specifiche.

Alcune di queste possono essere eliminate adottando la configurazione di ripresa ad assi paralleli, la quale - in teoria, adottando particolari misure - permetterebbe di ottenere una riproduzione in scala della scena ripresa, ma presenta diverse limitazioni sotto altri aspetti.

In pratica, l'effettivo impatto delle distorsioni sulla qualità soggettiva della scena ripresa deve essere ancora valutato a fondo e messo a confronto con altri elementi di giudizio derivanti dal processo di produzione.

Inoltre, è noto che le eventuali difficoltà di visualizzazione stereoscopica si attenuano man mano che l'utente si abitua a questo tipo di visualizzazione, e quindi il giudizio medio sulla qualità soggettiva dei sistemi stereoscopici è destinato a modificarsi - migliorando - nel tempo.

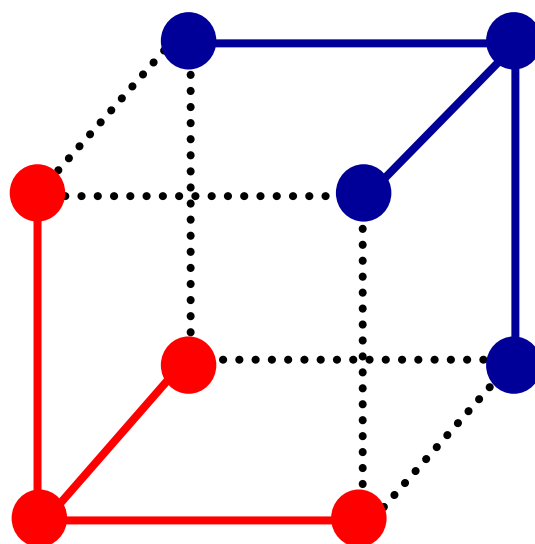
Tuttavia, alcune distorsioni non possono essere eliminate del tutto, in particolare quelle causate dal processo di visione umana e, probabilmente, anche quelle dovute alle tecniche di visualizzazione stereoscopica.

Bibliografia

1. M. Muratori: "La ripresa stereoscopica", Elettronica e Telecomunicazioni, n. 4, Dicembre 2007
2. H. Yamanoue, M. Okui, F. Okano: "Geometrical Analysis of Puppet-Theater and Coardboard Effects in Stereoscopic HDTV Images",IEEE Transactions on Circuits and Systems for Video Technology, Vol. 16, No. 6, June 2006
3. B.D. Diner: "A New definition of Ortostereopsis for 3-D Television", IEEE International Conference on Systems, Man, and Cybernetics, Conference Proceedings, Pages: 1053 – 1058, vol.2, 1991
4. A. Woods, T. Docherty, R. Koch: "Image Distortions in Stereoscopic Video Systems ", Proceedings of the SPIE, vol. 1915, San Jose, CA, Feb. 1993

Che cosa è, come funziona

Rivelazione, Correzione e Mascheramento degli Errori

Parte I

Marzio **Barbero**,
Natasha **Shpuza**

Sommario
Con questa serie di schede si vogliono fornire informazioni di base utili alla comprensione delle tecniche e ai sistemi digitali oggi ampiamente diffusi in campo televisivo. Questa prima parte è introduttiva alle tecniche utilizzati per la protezione dagli errori introdotti dal canale ed è focalizzata soprattutto sui codici a blocco e sulle applicazioni dei codici Reed Solomon in campo televisivo, in particolare per gli standard televisivi DVB di prima generazione e di videoregistrazione magnetica e ottica. Successivamente sarà pubblicata una seconda parte dedicata ai codici convoluzionali e alle tecniche utilizzate, ad esempio, negli standard DVB di seconda *generazione.*

1. Correggere gli errori per garantire la qualità

Il successo dei sistemi di comunicazione e memorizzazione digitali dei segnali audio e video è essenzialmente dovuto all'impiego in tali campi delle stesse tecnologie che sono state sviluppate per i calcolatori elettronici, e che, grazie alla loro evoluzione, consentono oggi di elaborare con velocità adeguate gli elevati flussi di dati associati alle informazioni audiovisive.

Le informazioni audio e video codificate secondo gli standard digitali [1,2] sono di tipo binario e ciò ha reso economicamente molto vantaggioso lo

sviluppo degli apparati digitali rispetto a quelli che operavano in ambito analogico. Infatti i segnali binari presentano solo due stati, o valori, in termini di corrente, tensione o magnetizzazione.

La trasformazione dei segnali video e audio in dati numerici consente inoltre di utilizzare tecniche atte a contrastare il degradamento che subisce l'informazione quando i segnali transitano sul canale hertziano o su cavo oppure vengono memorizzati. I canali, infatti, possono introdurre una serie di disturbi (rumore, interferenze, distorsioni, echi) che, combinandosi con il segnale utile, causano una non corretta identificazione dell'informazione in ricezione o in lettura.

Il flusso dati in ricezione non coincide quindi con quello trasmesso, ma a causa del canale, è caratterizzato da un certo tasso di errore (BER). Nel caso in cui i segnali audio e video vengano codificati utilizzando tecniche per la riduzione della ridondanza [3,4] è essenziale che praticamente tutti i dati trasmessi o memorizzati pervengano al decodificatore inalterati: solo in tal caso l'informazione originaria può essere riprodotta fedelmente, e senza degradamenti percettibili della qualità.

Tale condizione di "quasi assenza di errori" è quantizzabile, in termine di BER, con valori estremamente bassi, dell'ordine 10^{-10} (corrispondente ad esempio, nel caso di un flusso di 5 Mbit/s, a circa 2 bit errati ogni ora): solo con tassi di errore di quest'ordine di grandezza le immagini o l'audio codificate con le moderne tecniche di compressione possono essere riprodotte senza che i difetti introdotti dal canale siano considerati fastidiosi dall'utente, grazie alla bassissima frequenza con cui si presentano.

Una probabilità di errore così bassa può essere fornita, fra i canali normalmente utilizzati, solo dalla fibra ottica. La maggior parte degli altri canali è caratterizzata da BER superiori di diversi ordini di grandezza (arrivando anche a valori di $10^{-3} \sim 10^{-4}$): i sistemi digitali sarebbero poco competitivi, se non fossero stati sviluppati sistemi atti a rilevare e correggere gli errori.

Acronimi e sigle	
ATM	Asyncronous Transfer Mode
ATSC	Advanced Television Systems Committee (www.atsc.org)
BCH	Bose, Ray-Chauduri, Hocquenghem (codice di...)
BER	Bit Error Rate
BIS	Burst Indication Subcode
CD	Compact Disc
DSP	Digital Signal Processor
DVB	Digital Video Broadcasting, (www.dvb.org)
DVD	Digital Versatile Disc (www.dvdforum.org)
ECC	Error Correcting Code
FEC	Forward Error Correction
HDTV	High Definition TeleVision
LDC	Long Distance Code
LSI	Large Scale Integration
MDS	Maximum Distance Separable
PDM	Professional Disc Media
RAID	Redundand Array of Inexpensive Disks *oppure* Redundand Array of Independent Disks
RS	Reed Solomon (codice di...)

E' quindi stata fondamentale l'evoluzione, nel corso degli ultimi due decenni, delle tecniche di correzione basata sui codici; tecniche che hanno potuto trovare una applicazione pratica grazie all'incremento in termini di velocità e di capacità di memorizzazione di cui dispongono le attuali unità di elaborazione (microprocessori e DSP), in grado di effettuare in tempo reale le numerose e, a volte, complesse operazioni necessarie per la decodifica.

E’ quindi la progettazione di sistemi che integrano in modo ottimale le tecniche di compressione delle informazioni audio e video a quelle di protezione dagli errori e di modulazione, alla base dello spettacolare progresso tecnico della diffusione di immagini e suoni: aumento della densità di memorizzazione su supporti magnetici e ottici, aumento dell’efficienza spettrale delle comunicazioni, aumento dell’area di copertura dei servizi radiotelevisivi, diminuizione della potenza in trasmissione, via satellite e terrestre.

Si intende qui offrire una panoramica, a partire dalle prime ad essere introdotte in campo televisivo alla fine degli anni ‘70, delle tecniche per la rivelazione e correzione degli errori utilizzate nei sistemi per la diffusione e memorizzazione dei segnali video e audio digitali, pur evitando di approfondire le teorie alle basi dei codici, teorie che spesso implicano l’impiego di strumenti matematici complessi.

2. La rivelazione degli errori: il bit di parità

Il codice più semplice adottato per la rivelazione degli errori (*error detection*) introdotti da un canale di comunicazione è costituito dall’aggiunta di un solo bit di ridondanza per ciascuna delle parole in cui sono organizzati i dati trasmessi.

Il bit di parità, pari o dispari, consente di rilevare la presenza di errori, ma non consente di correggerli (figura 1).

E’ utile in presenza di probabilità di errore molto bassa, per verificare che il canale non presenti problemi di trasmissione, o nel caso in cui, avendo a disposizione un canale bidirezionale, sia possibile richiedere la ritrasmissione della parola errata. E’ efficace anche nel caso in cui l’informazione venga ritrasmessa ciclicamente, come avviene nel caso del servizio Televideo (figura 2) .

A	B	C
0	0	0
0	1	1
1	0	1
1	1	0

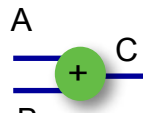



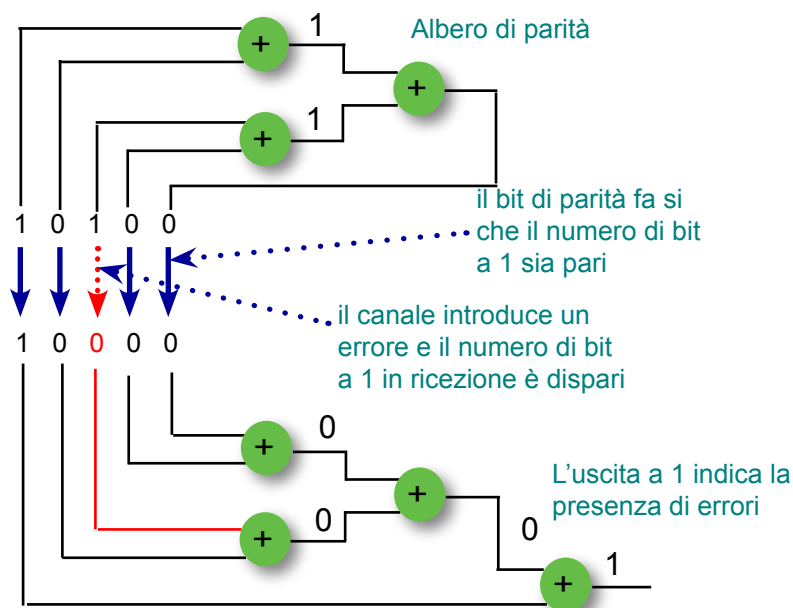
Fig. 1 - Vi sono due tipi di parità: il bit di parità pari e il bit di parità dispari.
Nel caso di bit parità pari, come in questo esempio basato su una parola di 4 bit di informazione, si aggiunge un bit 1 nel caso in cui la parola abbia un numero dispari di bit 1 oppure viene aggiunto uno 0 se tale numero è pari. Ciò si ottiene mediante semplici operazioni di somma binaria (XOR). Nell’esempio si trasmettono 5 bit. In ricezione la presenza, ma non la posizione, di un singolo bit errato è rivelata con modalità altrettanto semplici.
Nel caso di bit di parità dispari si fa sì che il numero di bit 1 trasmessi sia sempre dispari.

Una riga di cancellazione di quadro del segnale televisivo vista all'oscilloscopio: si vedono: il sincronismo di riga, il burst colore e i dati teletext.

Immagine di prova Televideo (Monoscopio Televideo) contiene tutti i caratteri di controllo e quelli visualizzabili.

In presenza di errori, essi sono facilmente identificabili nella pagina di test, che viene riacquisita ad ogni ciclo.

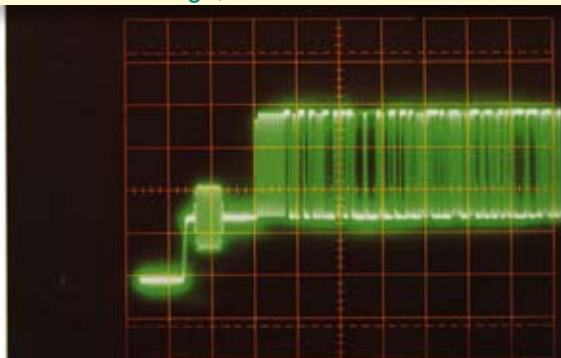



Fig. 2 - Nel sistema teletext, utilizzato per il servizio Televideo in Italia, i dati (pagine di 24 righe di 40 caratteri ciascuna) sono trasmessi sulle righe di cancellazione di quadro del segnale video.
I caratteri di testo o di controllo sono codificati come parole da 8 bit (byte o ottetto) di cui 7 bit sono di informazione e uno è il bit di parità.
Grazie al bit di parità è possibile individuare la presenza di un errore nel carattere ricevuto, in tal caso nella memoria di pagina viene memorizzato un indicatore (*flag*) e in visualizzazione lasciata vuota la posizione corrispondente.
In questo modo, nella successiva ricezione della stessa pagina (la trasmissione delle pagine televideo è ciclica), a meno che venga nuovamente rivelata la presenza di errore, il carattere può essere memorizzato e visualizzato. Anche nei casi in cui il canale sia affetto da una elevata probabilità di errore, il bit di parità risulta efficace come sistema per la rivelazione e correzione degli errori, grazie alla ridondanza insita nella ritrasmissione ciclica dell'informazione.
(foto tratte da [5]).

3. Codici a blocco e distanza di Hamming

In un codice a blocco i simboli (normalmente simboli binari) da codificare sono organizzati in parole di lunghezza finita prima di essere avviati al canale. Lo scopo della codifica è quello di consentire la localizzazione e la correzione di uno o più simboli, fra quelli costituenti la parola, che eventualmente pervengano errati al decodificatore.

Alla singola parola costituita da k simboli si aggiunge un numero predefinito di simboli in modo che la lunghezza totale della parola trasmessa, o memorizzata, diventi di n simboli. L'algoritmo di codifica aggiunge dunque n-k simboli di controllo, che costituiscono la ridondanza necessaria a permettere la localizzazione e correzione degli errori introdotti sul canale.

Nel caso in cui i simboli siano binari, cioè bit che possono assumere solo i valori 0 e 1, le combinazioni che possono essere generate dal codificatore sono 2^k, mentre le combinazioni possibili, e che possono pervenire al decodificatore, sono tutte le 2^n combinazioni degli n bit.

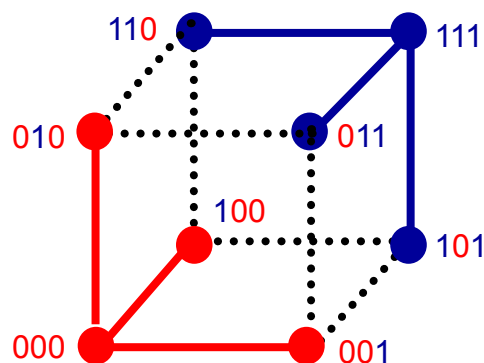
Un codice a blocco molto semplice è quello riprodotto in figura 3: in questo caso la decodifica è possibile grazie alla ripetizione del bit trasmesso e all'adozione, in fase di decodifica, di un criterio maggioritario, di massima verosimiglianza.

La stessa figura illustra il concetto di distanza (distanza di Hamming[Nota 1]) fra due parole: è il numero di posizioni per cui i simboli corrispondenti differiscono fra loro.

La distanza fra due parole rappresenta il numero di "errori" che trasformano una parola nell'altra. Se si calcolano i valori delle distanze fra tutte le coppie di parole codice che la sorgente può emettere e si individua il valore minimo, questo è detto distanza minima d_{min} e si può dimostrare che il numero e di errori correggibili in una parola è pari a $e = (d_{min}-1)/2$.

Un codice a blocco è quindi caratterizzato dai tre valori (n, k, d_{min}), la ridondanza è legata al numero di simboli aggiunti per la protezione, cioè n-k.

Nota 1 - Richard Hamming: nato a Chicago nel 1915 e morto a Monterey (CA) nel 1998.
Nel 1945 entrò a far parte del progetto Manhattan, il progetto per realizzare la bomba atomica, a Los Alamos.
Nel 1946 iniziò l'attività nei Bell Laboratories dove restò fino al 1976. Nel 1950 pubblicò l'articolo fondamentale sui codici per la rivelazione e la correzione degli errori [6].

Fig. 3 - In un codice con ripetizione il bit trasmesso è ripetuto n volte.
Ad esempio con un fattore di ripezione 3, le combinazioni trasmesse sono 000 e 111, in corrispondenza ai due simboli *0* e *1*, e si ottiene un codice a blocco (n=3, k=1).
In ricezione tutte le combinazioni di tre bit sono possibili, ma, assumendo che il canale non introduca più di un errore in una parola, si sceglie un criterio di massima verosimiglianza e quindi si decide che il simbolo trasmesso è quello meno distante da quello ricevuto.
La distanza d (detta distanza di Hamming) fra due parole è il numero di posizioni per cui i simboli corrispondenti differiscono fra loro.

In questo cubo ad ogni vertice corrisponde una delle 8 possibili parole da 3 bit. I vertici rossi hanno distanza ≤1 dalla parola 000, analogamente i vertici blu hanno distanza ≤1 dalla parola 111.

Non è possibile avere un codice che comprenda un elevato numero di parole di codice (valore di *k* alto), che corregga molti errori (d_{min} alto) e le cui parole di codice siano corte (*n* basso): occorre individuare il miglior compromesso fra i suddetti tre parametri.

Le tecniche di correzione degli errori che consentono l'identificazione diretta degli errori, senza richiedere la trasmissione ciclica o la ritrasmissione dell'informazione, sono definiti FEC.

4. Codice di Hamming

Un codice a blocco semplice è quello che Hamming ha proposto nel 1950 ed è utilizzato, ad esempio, nel teletext per identificare il numero di magazzino (in pratica la cifra delle centinaia che identifica la pagina Televideo), il numero di riga della pagina e le due cifre relative a decine e unità del numero di pagina.

Tali informazioni sono organizzate in gruppi di 4 bit e tali bit sono ritenuti molto importanti, perché associati alla corretta identificazione di pagina e riga. Una loro erronea decodifica implica la visualizzazione di una pagina non voluta o l'incorretta disposizione delle righe all'interno della pagina: in entrambi i casi il malfunzionamento è evidente all'utente, almeno per tutto il periodo intercorrente prima della ritrasmissione ciclica della pagina selezionata. Le informazioni associate ai caratteri visualizzabili sono invece

$$\begin{vmatrix} 1 & 0 & 0 & 1 \end{vmatrix} \begin{vmatrix} 1&0&0&0&0&1&1 \\ 0&1&0&0&1&0&1 \\ 0&0&1&0&1&1&0 \\ 0&0&0&1&1&1&1 \end{vmatrix} = \begin{vmatrix} 1&0&0&1&1&0&0 \end{vmatrix}$$

$$\begin{vmatrix} 0&0&0&1&0&0&0 \end{vmatrix}$$

$$\begin{vmatrix} 1&0&0&0&1&0&0 \end{vmatrix} \begin{vmatrix} 0&1&1 \\ 1&0&1 \\ 1&1&0 \\ 1&1&1 \\ 1&0&0 \\ 0&1&0 \\ 0&0&1 \end{vmatrix} = \begin{vmatrix} 1&1&1 \end{vmatrix}$$

in trasmissione si calcola: $x = u\,G$
il canale può aggiungere errori: $y = x + e$
in ricezione si calcola la sindrome: $s = y\,H^T$

Fig. 4 - Esempio di funzionamento del codice di Hamming (7,4).
A partire dai 4 bit generati dalla sorgente si calcolano 3 bit di ridondanza moltiplicando il vettore *u* di 4 bit (nell'esempio 1001) per la matrice *G* che consente di generare la parola di codice. Il calcolo consiste nel sommare (XOR) le righe di *G* corrispondenti ai bit 1 della parola (vettore) *u:* in questo esempio la prima e la quarta riga, in carattere corsivo in figura per facilitarne la comprensione.
La parola trasmessa risuta cosi essere il vettore *x* di sette bit.
Nell'esempio si suppone che il canale crei un errore sul quarto bit, che da 1 diventa quindi 0. La parola ricevuta è *y*, che differisce da quella trasmessa *x* per il bit in quarta posizione.
In ricezione si calcola un vettore detto sindrome (*s*), sommando le righe della matrice H^T (*H* è la matrice di controllo delle parità i cui elementi sono ricavati a partire dalla matrice generatrice *G*) corrispondenti agli 1 di *y* (in figura gli elementi delle righe di H^T che, sommati, danno origine alla sidrome, sono in carattere corsivo).
Se *s* è formato da tutti bit 0, si conclude che la parola ricevuta non è affetta da errore; negli altri casi, il valore di *s* indica quale è la posizione del bit errato.
In questo esempio *s*=[1 1 1] che corrisponde alla quarta riga di H^T, ovvero al quarto bit, che quindi viene riportato ad 1, correggendo l'errore introdotto dal canale.

ritenute meno importanti e, come si è visto, è ritenuto sufficiente l'uso del bit di parità per la rivelazione degli errori, la loro correzione è affidata alla ridondanza insita nella ritrasmissione ciclica delle pagine.

Questi gruppi di 4 bit sono protetti mediante un codice di Hamming (7,4,d_{min}=3), in grado di correggere errori singoli grazie a 3 bit di ridondanza. I codici di Hamming sono codici lineari e possono essere calcolati mediante l'uso di matrici (figura 4).

Si osservi che i 3 bit di ridondanza consentono di realizzare 2^3=8 configurazioni (denominate sindromi) di cui 7 sono utilizzate in ricezione per individuare quale dei 7 bit è affetto da errore e quindi correggerlo.

Il valore *n-k* è strettamente legato al numero di sindromi, e di conseguenza al numero di posizioni dell'eventuale errore identificabile: ad esempio il codice di Hamming (15,11) può utilizzare i 4 bit di ridondanza per individuare 16 sindromi, di cui quelle non nulle sono sufficienti a identificare la posizione dell'errore eventualmente presente in uno degli *n*=15 bit.

I codici correggono gli errori, in questo caso errori singoli, indipendentemente dal fatto che ne siano affetti i bit di informazione o quelli di ridondanza.

5. Effetti del superamento della capacità di correzione del codice

Nel caso in cui il numero di errori superi la capacità correttiva del codice, in ricezione, anziché recuperare l'informazione corretta, vengono introdotti ulteriori errori (figura 5). Se la capacità di correzione è insufficiente si ha una propagazione degli errori, in ricezione si passa rapidamente da una condizione ottimale (tutti gli errori sono recuperati e l'informazione risulta integra) ad una condizione di non funzionamento.

Questo fenomeno spiega le caratteristiche dei sistemi digitali confrontati con quelli analogici: i sistemi digitali sono caratterizzati da un rapido degradamento delle condizioni di servizio qualora si abbia un superamento della capacità di correzione degli errori, a differenza dei sistemi analogici che presentano un degradamento graduale delle prestazioni.

$$\begin{vmatrix} 1&0&0&1 \end{vmatrix} \begin{vmatrix} 1&0&0&0&0&1&1 \\ 0&1&0&0&1&0&1 \\ 0&0&1&0&1&1&0 \\ 0&0&0&1&1&1&1 \end{vmatrix} = \begin{vmatrix} 1&0&0&1&1&0&0 \end{vmatrix}$$

$$\begin{vmatrix} 0&0&1&1&0&0&0 \end{vmatrix}$$

$$\begin{vmatrix} 1&0&1&0&1&0&0 \end{vmatrix} \begin{vmatrix} 0&1&1 \\ 1&0&1 \\ 1&1&0 \\ 1&1&1 \\ 1&0&0 \\ 0&1&0 \\ 0&0&1 \end{vmatrix} = \begin{vmatrix} 0&0&1 \end{vmatrix}$$

Fig. 5 - Esempio di superamento della capacità di correzione del codice.
In questo caso il canale introduce due errori, nell'esempio sul terzo e quarto bit.
La sindrome determinata in ricezione in questo caso fornisce una indicazione errata e il settimo bit viene considerato errato, il decodificatore quindi non corregge i due bit errati, ma introduce un ulteriore errore che fa si che la parola decodificata diventi [1 0 1 0 1 0 1].

Occorre quindi progettare il sistema di correzione tenendo presenti le caratteristiche del canale e di criticità delle informazioni trasmesse, scegliendo il miglior compromesso fra ridondanza aggiunta e capacità di correzione o rilevazione degli errori.

Tornando all'esempio del sistema teletext, è ritenuto importante non solo proteggere gli indirizzi di pagina e riga mediante un codice, ma si è anche voluta limitare la probabilità di visualizzare una pagina diversa da quella selezionata a causa della presenza di errori in numero superiori all'errore singolo. Pertanto, ai 7 bit del codice (7,4,d_{min}=3) è stato aggiunto un ulteriore bit di parità e la distanza minima passa a d_{min}= 4: sono così rivelati gli errori di ordine pari, in particolare gli errori doppi.

In ogni caso gli errori di ordine dispari superiore a 1 (3, 5 e 7 errori) non possono essere rivelati e quindi in presenza di un così elevato numero di errori la capacità di correzione e rivelazione del codice viene superata.

I sistemi di correzione degli errori si basano sull'assunto che il BER sia basso e che gli errori siano distribuiti in modo uniforme all'interno del messaggio ricevuto, per evitare che all'interno del singolo blocco non si abbia, statisticamente, un numero di errori superiore a quelli correggibili.

6. Burst di errori e interleaving

Molto spesso gli errori vengono introdotti dal canale come salve o raffiche di errori (*error burst*). Ad esempio, nel caso della registrazione magnetica su nastro, la causa di errori può essere la mancanza di ossido o la presenza di particelle di polvere e quindi è interessata un'intera area del supporto e i bit ad essa corrispondente. Nel caso della diffusione, l'errore può interessare un simbolo demodulato, a cui corrispondono in genere più bit, oppure una temporanea riduzione della potenza del segnale ricevuto (*fading*) può comportare la mancata demodulazione di un certo numero di simboli.

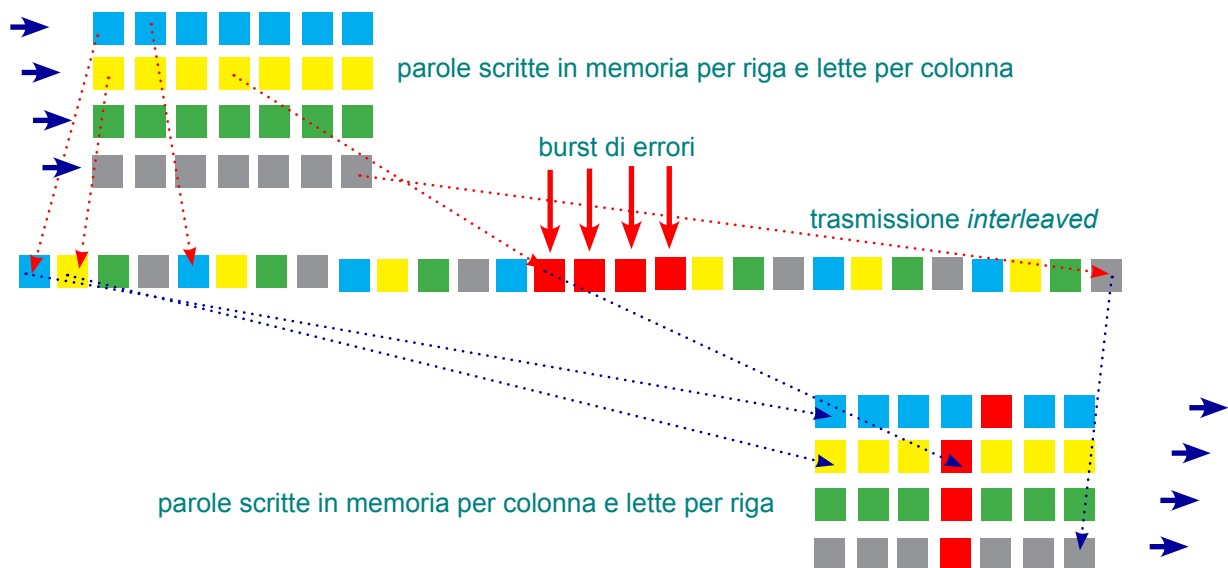



Fig. 6 - Esempio di interleaving: 4 parole di codice da 7 bit, ad esempio il codice di Hamming (7,4) in grado di correggere l'errore singolo, sono memorizzate e successivamente trasmesse sequenzialmente intercalando i bit relativi alle singole parole in modo che bit della stessa parola non siano consecutivi nel flusso trasmesso. Un eventuale burst di errori (in questo caso lungo 4) risulta così distribuito all'interno della memoria in ricezione: in questo esempio l'adozione dell'interleaving consente di recuperare tutta l'informazione, sfruttando appieno la capacità correttiva del codice. Ovviamente, sempre nell'esempio, se il burst fosse stato di lunghezza superiore, il codice non sarebbe stato in grado di correggere gli ulteriori errori.

Per ottenere una distribuzione uniforme degli errori e quindi limitare il più possibile gli effetti dei burst utilizza l'*interleaving* (figura 6), ovvero si intercalano le parole costituenti un blocco sufficientemente ampio di dati.

Questa tecnica non richiede un aumento della ridondanza, ma, grazie ad opportune memorie in registrazione e riproduzione, che però introducono un ritardo proporzionale alla lunghezza del codice a blocco e del fattore di *interleaving*, permettono di sfruttare al meglio le capacità correttive del codice.

L'*interleaving* a blocchi descritto in figura 6 è utilizzato in applicazioni dove è di scarsa importanza il ritardo introdotto per la scrittura nella memoria, quali i sistemi di videoregistrazione. D'altro canto una struttura regolare di intercalamento dei campioni può far si che gli eventuali errori non corretti creino una struttura anch'essa regolare di elementi di immagine errati, percepibile dal sistema psicovisivo. Per tale ragione in genere nel caso della videoregistrazione si utilizza una struttura non regolare nella memorizzazione dei campioni, in modo da diminuire la visibilità degli errori residui. Questa tecnica di *interleaving* viene denominata *shuffling* (in inglese, l'azione di mescolare le carte di un mazzo).

Nel caso dei sistemi di diffusione, dove è opportuno minimizzare i ritardi, in particolare per le trasmissioni in diretta, si utilizza uno schema differente, denominato *cross-interleaving,* che verrà descritto successivamente, nella seconda parte dell'articolo.

7. Codifica Reed-Solomon

Un codice a blocco che ha trovato ampio impiego nel campo delle telecomunicazioni, ed in particolare per i sistemi per la trasmissione e diffusione TV digitale DVB e ATSC, è denominato RS (dal nome di Irving Reed e Gustave Solomon che lo proposero nel 1960 [7]). Questo codice ciclico è un caso particolare dei codici BCH.

Il codice non è in teoria utilizzabile nel caso di simboli binari, ma questa limitazione si supera nella pratica organizzando i bit sotto forma di parole costituite da un certo numero di bit, in questo modo si possono considerare tali parole i simboli su cui applicare il codice RS.

7.1 Il più semplice...

Ad esempio il codice RS può essere costruito utilizzando simboli costituiti da 3 bit, sono quindi disponibili 2^3 simboli differenti; la lunghezza di questo codice a blocco è pari a $n=2^3-1=7$ simboli.

Di questi 7 simboli, 5 sono i dati utili e 2 sono i simboli di ridondanza, che indichiamo con P e Q. Non si riporta qui la trattazione matematica che permette di calcolare, mediante un sistema di due equazioni, due sindromi denominate *locator* e *corrector* che, quando non sono nulle, sono utilizzate rispettivamente una per localizzare la posizione dell'errore singolo (del simbolo errato) e l'altra per correggerlo. Questo codice consente quindi di correggere fino a 3 bit errati, se sono parte dello stesso simbolo.

Un vantaggio dei codici RS è la possibilità di dimezzare la ridondanza richiesta se, nel caso in cui non serva la funzione "localizzazione" perché è già nota la posizione degli errori, si utilizza solo la funzione "correzione". Gli errori di cui è nota la posizione sono generalmente denominati *erasure* e questa tecnica trova applicazione nella registrazione su supporto magnetico o ottico i cui dati posso risultare "cancellati" (ad esempio per mancanza di ossido o per la presenza di particelle di polvere tra il traferro della testina e il nastro).

Ad esempio, il codice RS con simboli a 3 bit può correggere un simbolo errato all'interno del blocco da 7 simboli utilizzando due simboli di ridondanza, ma se è utilizzato nell'ambito di uno schema che preveda le segnalazioni delle *erasure* è sufficiente un solo simbolo di ridondanza (e quindi 6 sono quelli di informazione) per la correzione di quello di cui è già nota la posizione. In alternativa si possono utilizzare

due simboli di ridondanza per correggere fino a due *erasure*.

7.2 ... e quello più usato

Al crescere del valore di *n* e di *k* aumenta il numero di equazioni che devono essere risolte dal decodificatore per localizzare gli errori. L'introduzione quindi di schemi di correzione così sofisticati, ma anche complessi, è stato quindi possibile solo ricorrendo a circuiti LSI. Molti degli standard di comunicazione attuali adottano i codici RS e quindi i produttori di DSP mettono oggi a disposizione librerie per consentire l'integrazione nei sistemi di codici RS con lunghezze con simboli da 3 a 12 bit, consentendo in questo ultimo caso di operare con blocchi fino a 2^{12}-1=4095 simboli.

Quello più diffuso è quello che adotta simboli costituiti da byte, organizzazione tipica dei dati in campo informatico, cioè da simboli che possono assumere 2^8 valori differenti. Scegliendo simboli di 8 bit, la lunghezza ottimale del blocco è $n= 2^8 -1=255$ simboli.

In funzione dell'applicazione, delle caratteristiche del canale e delle strategie di protezione si può scegliere il numero più opportuno di coppie di simboli utilizzati come *locator* e *corrector* o assegnare l'intera ridondanza alla correzione delle *erasure*.

La dimensione del blocco, 255 byte, è tale da rendere il codice particolarmente efficace nel caso in cui il canale sia caratterizzato dalla presenza di *burst* di errori molto lunghi. Il codice RS(255,239), in cui sono 16 i byte di ridondanza, fu utilizzato nella sperimentazione di trasmissione punto-punto via satellite in HDTV effettuata in occasione dei campionati mondiali di calcio Italia '90 e fu in seguito adottato nello standard ITU-T J81, per le reti di contributo televisive.

Il codice è in grado di correggere fino a $(n-k)/2$ simboli errati: e con k =239 si hanno 16 simboli di ridondanza, sufficienti a localizzare e correggere fino a 8 byte errati nel blocco, mentre se lo schema di codifica prevede che la localizzazione sia fatta a monte, la capacità di correzione sale a 16 *erasure*. E' pertanto possibile correggere anche burst molto lunghi, fino a 8 o 16 byte.

Questo codice, in una versione accorciata, è utilizzata nei sistemi di diffusione televisiva DVB di prima generazione (DVB-S, DVB-C, DVB-T).

7.3 Codici accorciati

Si è visto che il parametro *n*, cioè la lunghezza dei blocchi, dipende dal numero di bit associati a ciascun simbolo. D'altro canto a volte è conveniente organizzare il flusso binario in funzione delle caratteristiche del messaggio emesso dalla sorgente o della struttura di sincronizzazione del multiplex e ciò restringe la scelta dei valori di *k* e *n*.

Una tecnica che consente di assegnare un numero prefissato *k* di simboli di informazione a ciascun blocco è quella di accorciare il codice prescelto (figura 7), anche se ciò avviene a spese di una diminuizione del tasso di informazione, cioè il rapporto *k/n*, che caratterizza il codice.

I sistemi DVB utilizzano il codice accorciato RS(204,188)[Nota 2] ricavato da RS(255,239) limitando a 204 il numero di byte utili trasmessi in ogni pacchetto, mentre il numero di byte di ridondanza rimane 16, mantenendo la capacità di correggere 8 byte oppure 16 nel caso di *erasure*.

Nota 2 - La scelta di *k*=188 fu determinata dal requisito, ritenuto importante agli inizi degli anni '90, di facilitare il trasferimento delle informazioni audio e video compresse sui sistemi di ponti radio e di satelliti previsti per le telecomunicazioni, basati sullo standard ATM. Tale sistema, progettato per la commutazione e trasmissione telefonica , prevede che i dati siano organizzati in celle di 53 byte, di cui 48 byte di dati e 5 byte di intestazione. Il valore 188 (di cui i primi 4 costituiscono l'intestazione) fu scelto per facilitare la rimappatura dei dati utili codificati secondo gli standard MPEG nelle celle ATM.

8. Dai codici prodotto al picket code

I codici RS trovano impiego nei sistemi basati su nastro magnetico o disco ottico per la protezione dagli errori a *burst* che, a causa di imperfezioni del mezzo di registrazione o della presenza di materiale estraneo tra esso e il sistema di lettura, possono essere presenti nel flusso di dati riprodotti.

Normalmente tali sistemi usano codici prodotto ottenuti con una coppia di codici organizzati nello schema *outer* e *inner code* (figura 8).

Si tratta di codici accorciati derivati da Reed Solomon con *n*=255 e simboli costituiti da byte. Nel corso degli anni, al crescere delle capacità di elaborazione dei decodificatori, è stato possibile adottare anche nel campo della registrazione codici con valori più elevati di *k* e *n,* approssimandosi a 255, e migliorando conseguentemente le prestazioni del sistema ECC, pur riducendo contemporaneamente la ridondanza ad esso associata.

Nel 1982 fu definito il formato del CD che adottava come sistema ECC la coppia di codici RS(32,28) e RS(28,24). Il primo, il codice interno, è in grado di correggere fino a due simboli e di rivelare la presenza di un numero superiore di errori. Il fattore di *interleaving* utilizzato è pari a 28. Il codice esterno è in grado di correggere fino a 4 *erasure,* garantendo così l'elevata capacità di correggere *burst* molto lunghi, fino ad un massimo di 4000 bit, corrispondenti a circa 2,5 mm in termini di lunghezza della traccia del disco.

Nel 1987 fu normalizzato il D1, il primo registratore digitale per segnale televisivo, a componenti e per usi professionali. Adottava un nastro da 3/4" e l'ECC era basata sulla coppia RS(32,30) e RS(64,60).

Nel 1991, il D3, sistema di registrazione televisiva digitale in formato composito su nastro da 1/2", fu in grado di migliorare notevolmente la densità di registrazione sul nastro, anche grazie all'adozione di un ECC basato sulla coppia RS(166,158) e RS(84,76).

Fig. 7 - Esempio di codice accorciato: si suppone di utilizzare un codice RS i cui simboli siano associati a 3 bit, in questo caso $n=2^3-1=7$ e con k=5 si è in grado di localizzare e correggere 1 simbolo errato. Il codice RS(7,5) implica quindi un blocco di 21 bit di cui 15 utili e 6 di ridondanza. Si supponga di voler organizzare i blocchi in modo che i bit utili siano solo 12. In questo caso è sufficiente calcolare i due simboli di ridondanza a partire da un simbolo fittizio, costituito da tutti 0, e da 4 simboli utili: ovviamente il simbolo costituito da tutti 0 non viene trasmesso e il blocco effettivamente posto sul canale è costituito da 12 bit utili e 6 di ridondanza: si parla di codice accorciato RS(6,4). In ricezione si provvede a ricostruire il blocco RS(7,5), prima di identificare e correggere gli errori.
Può succedere che il superamento della capacità correttiva del codice porti a localizzare errori anche nei simboli fittizi, ma in questo caso il decodificatore non effettua ovviamente alcuna correzione, ma rivela il superamento della capacità correttiva del codice e la presenza nel blocco di errori non correggibili, perché non localizzabili, agli eventuali stadi di correzione o mascheramento successivi.

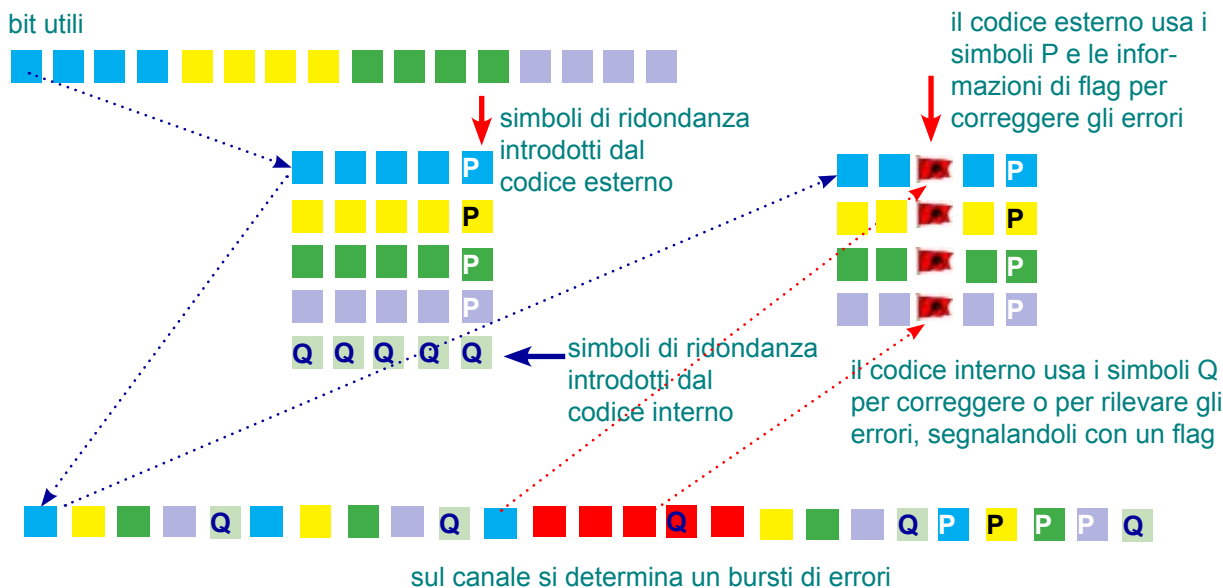



Fig. 8 - Esempio di schema di codice prodotto. Il flusso di dati utili è organizzato in blocchi, i simboli di protezione P sono calcolati in base ai dati disposti per righe; sucessivamente vengono calcolati i simboli Q, calcolati in base ai dati disposti per colonne.
Questo secondo codice è denominato codice interno (*inner code*) perché è applicato immediatamente prima del processo di trasmissione sul canale, mentre il codice che ha prodotto i simboli P è denominato codice esterno (*outer code*) perché è più lontano dal canale, nello schema codifica-canale-decodifica.
L'insieme dei dati vengono memorizzati o trasmessi sequenzialmente, sfruttando così la tecnica di *interleaving*. Nell'esempio si suppone che il canale introduca un burst di errori su cinque simboli consecutivi. In ricezione viene adottata la decodifica utilizzando i simboli Q che può consentire la localizzazione e correzione di errori. Nel caso in cui la capacità di correzione del codice interno sia superata, ma sia possibile rilevare la presenza di errori residui, non correggibili, questo evento è segnalato (con un *flag*) al secondo livello di correzione.
Nell'esempio si suppone che il codice interno sia in grado di correggere fino ad un simbolo errato e rivelare la presenza di errori in numero superiore, quindi corregge l'ultimo dei cinque simboli errati e rivela con un flag la presenza di errori nei quattro simboli di informazione che costituiscono la parola che precedente. A questo punto il codice esterno, che si basa sui simboli ridondanti P, può essere in grado di correggere gli errori residui.

Lo standard DVD, del 1995, utilizza una coppia RS(182,172), come *inner code*, e RS(208,192), come *outer* code.

Nel 2000 è stato definito il formato di videoregistrazione digitale professionale a componenti e con compressione MPEG-2 su nastro da 1/2" denominato D10 (noto anche come Betacam IMX). L'ECC è basato sulla coppia RS(162,150) e RS(64,54): considerando anche le informazioni di sincronismo e di indentificazione dei blocchi, la ridondanza è pari al 28% e consente la corretta riproduzione dell'informazione video anche in caso di abrasioni pari 1,16 mm per le tracce video. I dati audio sono protetti con la coppia RS(137,125) e RS(18,8), ovvero il 147% di ridondanza per consentire la corretta riproduzione in caso di abrasioni pari 1,12 mm per le tracce audio.

Nel caso del disco ottico PDM, utilizzato per il sistema di videoregistrazione professionale XDCAM, è stato realizzato uno schema ECC denominato *Picket Code.* Tale schema è basato su

una coppia di codici RS(62,30) e RS(248,216), denominati rispettivamente BIS e LDC. LDC è in grado di correggere *burst* di lunghezza molto elevata (tabella 1). I due codici sono intercalati in modo tale che vi sia un byte del codice BIS ogni 38 byte del codice LDC. Il codice BIS è molto robusto ed è in grado di correggere fino a 15 byte, che, grazie al fattore di *interleaving*, sono distribuiti su un blocco di dati lungo fisicamente 76,66 mm. La presenza di più errori consecutivi nei byte costituenti il BIS è utilizzato in fase di decodifica per rivelare la presenza e la posizione di eventuali burst di errori e ciò rende possibile sfruttare al meglio la capacità del codice LDC di correggere efficacemente le *erasure*, oltre agli errori isolati.

9. Un'incursione fra le schiere di dischi

Allo scopo di riassumere le tecniche di correzione degli errori precedentemente trattate, può essere utile analizzare una delle applicazioni più diffuse: la protezione delle informazioni memorizzate su hard-disk, per consentirne il recupero in caso di guasti.

Nel 1987, un gruppo di esperti di informatica dell'Università di California a Berkeley studiava metodi per realizzare schiere (*array*) di dischi (di basso costo, e quindi limitati in termini di capacità e tempi di accesso) che apparissero al computer ospite come un singolo disco, di grande capacità e con elevata velocità di lettura e scrittura.

Un primo metodo individuato consiste nel suddividere, a livello della scheda di controllo dei dischi, l'insieme di dati (considerato come una pagina) in tante strisce (*stripe*), ciascuna delle quali può quindi essere scritta (e letta) su uno degli *n* dischi che compongono la schiera. Poiché si può effettuare l'accesso in parallelo ad *n stripe,* il metodo permette di incrementare di un fattore prossimo a *n* la velocità accesso ai dati, ma il malfunzionamento o guasto di uno dei dischi può causare la perdita dell'intera informazione e la probabilità che ciò accada è proporzionale al numero di dischi, multipla della probabilità che si guasti un singolo disco.

Occorre quindi individuare delle tecniche per proteggere le informazioni e gli esperti individuarono 5 livelli di protezione che denominarono RAID, cioè schiera ridondante di dischi a basso costo.

Il livello precedentemente descritto, che non fornisce alcuna protezione, ma non richiede nessuna ridondanza, fu definito il livello RAID 0.

	DVD	PDM
Distanza tra le tracce (Track Pitch) [μm]	0,74	0,32
Lunghezza di un Byte registrato [μm]	2,13	0,96
Area occupata da 1 Byte sul disco [mm^2/Byte]	1,576 10^{-6}	0,3072 10^{-6}
Numero di Byte per mm^2	635 KB	3,255 MB
Numero di Byte al mm	469 Byte	1041 Byte
Burst di errori correggibili	2790 Byte	9982 Byte
Lunghezza del burst correggibile [mm]	5,95	9,58

Tab. 1 - Confronto tra la capacità correttiva per burst di errori nel caso di un disco DVD e PDM. Il PDM utilizza luce a frequenza più elevata (blu-violetta) rispetto al DVD (rossa), e protetto in un cartridge per ridurre il rischio di contaminazione della superficie da parte di particelle estranee.
I dati riportati ipotizzano che una particella di polvere di 1 mm^2 sia depositata sulla superficie dei due media a confronto. I dati coinvolti sono, per il PDM, circa 5 volte in più, a parità di superficie, e circa il doppio a parità di lunghezza della traccia.
E' stato quindi sviluppato un sistema ECC, denominato *Picket Code*, che si basa su una coppia di codici Reed Solomon, RS(62,30) e RS(248,216), particolarmente adatto a rilevare burst molto lunghi, ma che non sfrutta lo schema classico di *inner* e *outer code*,

Nel primo livello (RAID 1) i dati presenti su ciascun disco sono scritti anche su un secondo disco che quindi contiene, specularmente, tutta l'informazione (*mirroring*). Nel caso in cui uno dei dischi si guasti, può essere sostituito con uno nuovo, su cui si copiano i dati presenti sul disco ancora funzionante. Ovviamente occorre utilizzare il doppio dei dischi strettamente necessari e come minimo occorre avere due dischi. Dal punto dei sistemi di codifica per la protezione dagli errori, questo è un codice con ripetizione (ciascun bit è ripetuto due volte).

Il secondo livello (RAID 2) è basato sul codice di Hamming (7,4). Le informazioni sono divise in 4 *stripe* e scritte su 4 dischi, mentre altri 3 dischi contengono i 3 bit di parità calcolati in base ai 4 corrispondenti bit di informazione. Questo livello non ha mai trovato applicazione pratica, a causa del costo legato alla ridondanza, 7 dischi anziché 4.

I successivi due livelli (RAID 3 e RAID 4) organizzano i dati rispettivamente in byte o in blocchi,e li distribuiscono su k dischi; per ciascuno dei k bit presenti su tali dischi, calcolano un bit di parità che viene memorizzato su un ulteriore disco. Il numero totale di dischi è quindi $n=k+1$ e il numero minimo di dischi necessari è 3, due per i bit di informazione ed uno per i bit di parità. L'organizzazione dei dati in blocchi (RAID 4) è ovviamente più flessibile e può consentire tempi di accesso migliori rispetto all'organizzazione in byte (RAID 3).

Il livello RAID 5 è quello attualmente più utilizzato. I dati sono organizzati in blocchi, come nel caso del RAID 4, ma i blocchi di dati e di parità sono distribuiti uniformemente sugli k+1 dischi, anziché concentrare i blocchi di parità su un solo disco (figura 9). Questa struttura implica la stessa ridondanza rispetto ai due livelli precedenti e garantisce la stessa protezione in caso di guasto di un disco costituente la schiera, ma distribuisce uniformemente le operazioni di scrittura e lettura fra gli n dischi, migliorando le prestazioni dell'intero sistema. L'operazione di scrittura richiede mediamente più accessi al disco rispetto a quella di lettura poiché in genere, dopo l'operazione di scrittura, il *controller* rilegge il dato al fine di verificare l'assenza di errori, ed in caso di non coincidenza fra il dato scritto e quello riletto, effettua ulteriori tentativi.

I tre livelli (RAID 3, 4 e 5) consentono il recupero dell'informazione originaria nel caso di malfunzionamento o guasto di uno dei dischi. In genere il semplice bit di parità non permette la correzione dell'errore, solamente la rivelazione, ma in questo caso invece l'errore può anche essere corretto perché è nota anche la sua posizione: si sa quale degli n dischi è guasto. E' quindi

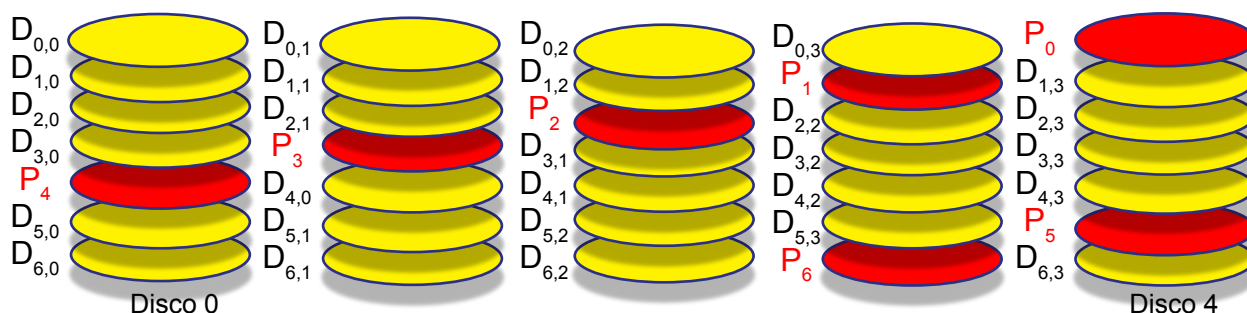



Fig. 9 - Semplice esempio di organizzazione dei dati nel caso del RAID 5, utilizzando 5 dischi: nella rappresentazione schematica in figura, ciascun blocco di dati è rappresentato da una ellisse, l'insieme di ellissi incolonnate rappresenta un disco.
Le parità P_0 ottenute a partire dai primi 4 blocchi di dati $D_{0,0}$...$D_{0,3}$ sono memorizzate nel quinto disco; P_1, ottenuto dai successivi 4 blocchi $D_{1,0}$...$D_{1,3}$ nel quarto disco ...

sufficiente sostituire al disco guasto un disco di scorta e il sistema provvede a rigenerare i dati perduti, calcolandoli a partire da quelli contenuti nei k=n-1 dischi funzionanti.

L'obiettivo originario, nel 1987, era l'uso di schiere di dischi di basso costo, ma il risultato è una struttura non necessariamente *Inexpensive*, cioè poco costosa, ma più affidabile, grazie alla possibilità di sostituire i singoli dischi: oggi il significato originario (*Inexpensive*) della I nell'acronimo RAID è stato modificato per evidenziare il fatto che i dischi sono indipendenti (*Independent*).

La protezione offerta dal livello RAID 5 è generalmente sufficiente, ma richiede che la sostituzione del disco guasto avvenga in tempi brevi, infatti il tempo che intercorre fra il verificarsi del guasto e la sostituzione è estremamente critico: un ulteriore guasto porta alla perdita di tutti i dati. Inoltre può essere a volte necessario disconnettere la schiera di dischi nel corso dell'operazione di sostituzione.

Al giorno d'oggi i server utilizzati per applicazioni video e audio, ad esempio per *streaming*, e in genere per web ed archivi, hanno dimensioni tali per cui richiedono schiere di dischi sempre più capaci. All'aumentare del numero di dischi che compongono il RAID aumentano sia la probabilità di guasto per il singolo disco che gli incovenienti causati dalla sospensione del servizio e pesanti sono i danni dovuti eventualmente al verificarsi di un guasto non recuperabile.

La probabilità di un difetto non recuperabile cresce con il numero di dischi che costituisce la schiera: nel corso della ricostruzione dei dati a seguito della sostituzione di un disco, il controller può rivelare un settore difettoso di uno dei dischi ritenuti funzionanti (evento con probabilità non trascurabile se i dischi sono numerosi) e i dati memorizzati in tale settore non possono più essere ricostruiti poiché manca la ridondanza, non più disponibile a causa della sostituzione del disco.

Sono stati pertanto sviluppati controller in grado di utilizzare un nuovo livello di protezione (RAID 6) basato sui codici MDS.

Tali codici sono un caso particolare del codice Reed Solomon con simboli costituiti da byte. Si suppone di utilizzare una schiera di 255 dischi, in modo da distribuire in dati sotto forma di parole di 255 simboli, di cui 253 sono le informazioni originali e 2 sono i simboli di ridondanza. Il codice è utilizzato per la caratteristica di poter correggere errori di cui è nota la posizione (*erasure*) e pertanto la possibilità di perdere informazione è limitata solo ai casi, estremamente poco probabili, in cui più di due dischi si guastino

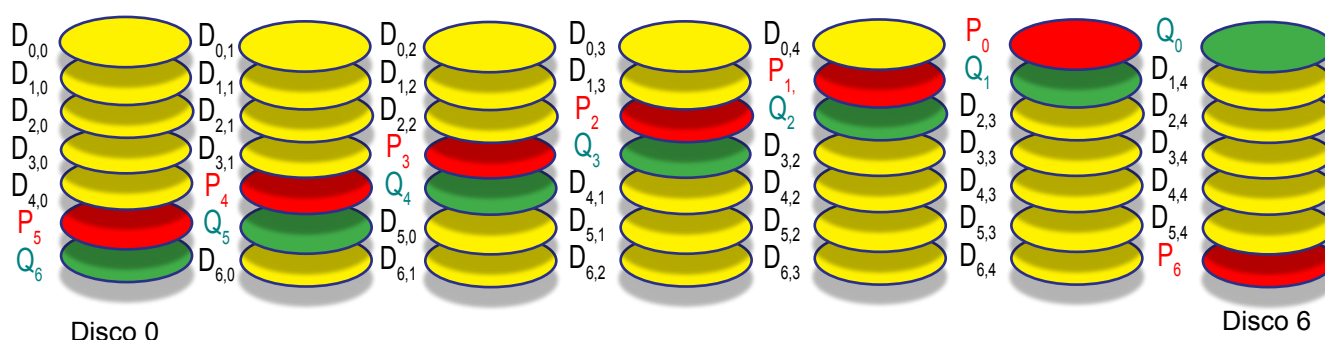



Fig. 10 - Semplice esempio di organizzazione dei dati nel caso del RAID 6, utilizzando 7 dischi (le colonne della tabella). Le parità P_0 e Q_0 ottenute a partire dai primi 5 blocchi di dati $D_{0,0}$...$D_{0,5}$ sono memorizzate rispettivamente nel sesto e settimo disco; P_1 e Q_1, ottenute dai successivi 5 blocchi $D_{1,0}$...$D_{1,4}$ nel quinto e sesto disco ...

contemporaneamente, oppure se l'informazione "cancellata" debba essere ricostruita a partire da più settori che risultino difettosi al momento della sostituzione del disco.

Ovviamente non è necessario che la schiera RAID 6 sia costituita da 255 dischi: si può utilizzare una versione accorciata del codice (figura 10), basta garantire la presenza di due dischi aggiuntivi, per i simboli di ridondanza, cioè che $n=k+2$.

Il vantaggio di utilizzare lo schema RAID 6, e quindi un codice più complesso come il Reed Solomon, rispetto a quello RAID 5, e quindi la semplice parità, è evidente nel caso in cui si considerino capacità complessive elevate ottenute con un grande numero di dischi e tenendo in considerazione la probabilità di trovare difetti latenti durante la fase di ricostruzione. Ad esempio con capacità complessive superiori a 5 TB ottenute con dischi da 320 GB, il passaggio da una schema RAID 5 ad uno RAID 6, può ridurre il tempo medio fra perdite di dati da valori dell'ordine di un mese a valori prossimi a 100 anni [8].

(1-continua)

Bibliografia

1. M. Barbero, N. Shpuza: "Le origini del video digitale (la rac. ITU-R BT.601)", Elettronica e Telecomunicazioni, aprile 2003
2. M. Barbero, N. Shpuza: "I format HDTV (le rac. ITU-R BT.709 e BT.1543)", Elettronica e Telecomunicazioni, aprile 2005
3. M. Barbero, N. Shpuza: "Uno standard pervasivo (MPEG-2 video)", Elettronica e Telecomunicazioni, aprile 2003
4. M. Barbero, N. Shpuza: "Advanced Video Coding (AVC - H.264)", Elettronica e Telecomunicazioni, aprile 2003
5. M. Cominetti, D. Tabone:"televideo: la telematica in ogni casa", Rai, Direzione Commerciale, 1987
6. R. W. Hamming: "Eror detecting and error correting codes", Bell Syst. Tech. J. 29:147-60,1950
7. 9. I.S. Reed, G. Solomon: "Polynomial codes over certain finite fields", J. Soc. Indust. Appl. Math, pp. 300-304, 1960
8. "Intelligent RAID 6 Theory. Overview and Implementation", Intel whitepaper, 2006